UNIVERSITY OF OSLO

Department of informatics

Audio conferencing

Study of User experience with the application of Activity theory

Master thesis

60 credits

Irena Daskalova

10.May 2010

Acknowledgements

This thesis presents my academic work for the Master Degree in “Information Systems” at the Department of Informatics, University of Oslo.

I would like to thank my advisors Alma Culen (internal), Peter Brooks (external) and Bjørn Hestnes (Telenor) for their guidance, advice and support during my fieldwork and during the writing of the thesis and for the harmonious atmosphere in the team.

I would also like to express my gratitude to Svein Heiestad from Telenor who made the web-based survey and to all the people from Telenor who made possible my fieldwork in the company and to those who participated as respondents in the interviews and the survey.

Additionally, my thanks to Jo Herstad, Sara Kalantari and Kim Anh Ti Vo from IFI, and my friends Radoslava Dimitrova and Velichka Gotskova with whom I piloted the survey.

Special thanks to my skillful sister Dimitrina for her goodness and willingness to support me in every stage of this work.

Content

ACKNOWLEDGEMENTS **3**

CONTENT **5**

LIST OF FIGURES AND TABLES **10**

1. INTRODUCTION **13**

1.1 OBJECTIVES AND MOTIVATION 14

1.2 RESEARCH QUESTIONS 17

2. LITERATURE REVIEW **20**

2.1 AUDIO CONFERENCING 20

2.1.1 Audio conference equipment *20*

CHALLENGES IN AUDIO CONFERENCING SESSIONS 21

2.2 21

2.2.1 Technical challenges *22*

2.2.2 Socio-cultural communication challenges *24*

2.2.3 Language challanges in distributed communication *28*

2.3 RELATIONSHIP BETWEEN QOS AND QOE 29

2.4 STUDY OF USER EXPERIENCE 31

3. THEORETICAL FRAMEWORK **32**

3.1 IDENTIFYING THREE MAIN PERSPECTIVES 32

3.2 SEARCHING FOR A SOCIO-TECHNICAL FRAMEWORK FOR THE ANALYSIS 34

3.3 ACTIVITY THEORY 36

3.3.2 Developing an analytical approach based on AT *37*

3.4 DISTRIBUTED COGNITION THEORY (DCOG) 42

3.5 ACTOR-NETWORK THEORY (ANT) 43

3.6 STRAUSS' THEORY 44

3.7 CONCLUSION 45

4. DESCRIPTION OF THE FIELD SETTING 47

4.1 TELENOR 47

4.2 AUDIO CONFERENCING SYSTEMS IN TELENOR 48

4.2.1 Audio conferencing systems using PSTN network 48

4.2.2 Internet applications 49

5. METHODOLOGY 52

5.1 ETHICAL ISSUES 52

5.2 ETHNOGRAPHY 53

5.3 ETHNOGRAPHY IN COMPUTER SCIENCES 54

5.4 RELATIONSHIP BETWEEN RESEARCH METHODOLOGY AND THEORETICAL FRAMEWORK 55

5.5 LIMITATIONS 56

6. INTERVIEW STUDY 58

6.1 AIM 58

6.2 PARTICIPANTS 58

6.3 METHOD 59

6.4 RESULTS AND DISCUSSION 61

6.4.1 Technical issues 62

6.4.2 Socio-cultural communication issues 63

6.4.3 Language issues 68

6.5 CONCLUSION 71

7. OBSERVATION STUDY 73

7.1 AIM OF THE STUDY 73

7.2 PARTICIPANTS 73

7.3 METHOD 74

7.4 RESULTS 74

7.4.1 Visibility who is in the meeting 75

7.4.2 Mute button 75

7.4.3 Background noise 75

7.4.4 Volume control 76

7.4.5 Accented speech 76

7.4.6 Text messages supporting audio conference 76

7.4.7 Identify and show who is speaking 76

7.4.8 Overlapping 77

7.5 DISCUSSION OF THE RESULTS 77

8. QUESTIONNAIRE STUDY 79

8.1 AIM OF THE QUESTIONNAIRE STUDY 79

8.2 METHOD – DESCRIBTION OF THE QUESTIONNAIRE STUDY 80

8.2.1 Design of the content of the survey 81

8.2.2 Design of the questions and the answers in the survey 83

8.2.3 Testing the questionnaire and piloting the survey 86

8.2.4 QuestBack 86

8.3 PARTICIPANTS 87

8.3.1 Target population and Samples 87

8.4 RESULTS AND DISCUSSION 88

8.4.1 Activity theory and activity system 88

8.4.2 Activity = Subject – Object interaction 90

8.4.3 Subject 96

8.4.4 Tool 102

8.4.5 Relationship Subject – Tool 103

8.4.6 Relationship Subject – Rules 106

8.4.7 Community 109

8.4.8 Outcome of activity 111

9. CONCLUSION 115

9.1 WHAT ARE THE TECHNICAL DIFFICULTIES AND WHAT KIND OF APPROACHES COULD BE OFFERED TO OVERCOME THEM? 116

9.2 WHAT ARE THE SOCIO-CULTURAL DIFFICULTIES AND WHAT KIND OF APPROACHES COULD BE OFFERED TO OVERCOME THEM? 117

9.3 WHAT ARE THE LINGUISTIC DIFFICULTIES AND WHAT KIND OF APPROACHES COULD BE OFFERED TO OVERCOME THEM? 117

9.4 CONTRIBUTIONS 118

9.4.1 AS's Matrix 118

9.4.2 Categorised statistical data for the current case 119

9.4.3 Guidelines for improving user experience: Meeting protocol / etiquette 119

10. FUTURE WORK 120

REFERENCES 121

APPENDIX A: GUIDELINES FOR IMPROVING USER EXPERIENCE IN AUDIO CONFERENCING SESSIONS: MEETING PROTOCOL / ETIQUETTE 1

Socio-cultural issues 1

Language issues 2

Technical issues (on users level) *3*

APPENDIX B: GENERAL CONCEPTS AND TERMINOLOGY **4**

APPENDIX C: NOTES TAKEN FROM THE SECOND OBSERVATION **6**

APPENDIX D: SURVEY QUESTIONNAIRE **7**

LANGUAGE 7

COMMUNICATION 8

Purpose and language *8*

Meeting Protocol *9*

Destination *9*

Language *18*

AUDIO CONFERENCING SYSTEM 18

CHALLENGES IN AUDIO CONFERENCING 20

Audio quality issues *22*

Socio-cultural issues *26*

FINAL QUESTIONS 41

List of figures and tables

FIGURE 1: CONFERENCE PHONE 20
FIGURE 2: EXPLANATORY DIAGRAM, MEDIA RICHNESS THEORY; SOURCE: WIKIPEDIA.ORG 26
FIGURE 3: THE THREE MAIN LEVELS OF THE WORK PROCESSES, ACCORDING TO HEEREN & LEWIS (1997) 27
FIGURE 4: RELATIONSHIP BETWEEN QOS AND QOE 30
FIGURE 5: ENGESTRÖM'S (1999) ACTIVITY SYSTEM MODEL 36
FIGURE 6: HIERARCHICAL STRUCTURE OF ACTIVITY 37
TABLE 1: MATRIX REPRESENTING THE RELATIONSHIP BETWEEN EACH TWO AS'S ELEMENTS 38
FIGURE 7: TELENOR WORLD LOCATIONS 2009, SOURCE WIKIPEDIA.ORG 47
FIGURE 8: CONFERENCE PHONE BY KONFTEL IN ONE OF THE AUDIO CONFERENCING ROOMS IN TELENOR R&I 49
FIGURE 9: LIKERT RATING SCALE 85
FIGURE 10: SEMANTIC SCALE; 1 IS EXTREMELY NOT IMPORTANT, 7 IS EXTREMELY IMPORTANT 85
FIGURE 11: HIERARCHICAL STRUCTURE OF ACTIVITY 90
FIGURE 12:"HOW OFTEN ARE YOU USING THE AUDIO CONFERENCING SYSTEMS AT TELENOR?" 91
FIGURE 13: THE NATURE OF THE PROBLEMS EXPERIENCED DURING AUDIO CONFERENCING 92
TABLE 2: BEHAVIOURAL PATTERNS REDUCING QOE IN AC 94
FIGURE 14: BEHAVIOURAL PATTERNS REDUCING QOE IN AC 95
TABLE 3: THE MEAN VALUES FOR AUDIO QUALITY, EFL, MEETING PROTOCOL FOR 12 COUNTRIES; 1 IS EXTREMELY LOW QUALITY, 7 IS EXTREMELY HIGH 98
FIGURE 15: AUDIO QUALITY FOR 12 COUNTRIES; 1 IS EXTREMELY LOW AUDIO QUALITY, 7 IS EXTREMELY HIGH AUDIO QUALITY 99
FIGURE 16: THE COMMUNICATION IN EFL WITH 12 COUNTRIES; 1 IS EXTREMELY LOW, 7 IS EXTREMELY HIGH .. 100
FIGURE 17: THE MEETING PROTOCOLS FOR 12 COUNTRIES 101
FIGURE 18: "WHAT EQUIPMENT DO YOU USE FOR AC? (HEADSET/ HANDSET/ LOUDSPEAKER/ MOBILE PHONE/ PC/ OTHER)" 102
FIGURE 19: HIERARHICAL MODEL OF AT 104
FIGURE 20: "WHAT KIND OF COMMUNICATION ARE YOU USING THE AUDIO CONFERENCING FOR?" 105
FIGURE 21: "WHAT IS THE PURPOSE OF THE AFTER-CALL MAIL?" 106
FIGURE 22:"DO YOU FOLLOW THE MEETING PROTOCOL?" 107
FIGURE 23:"WITH WHICH COUNTRIES HAVE YOU BEEN COMMUNICATING OR CURRENTLY COMMUNICATE WITH IN ENGLISH THROUGH AUDIO CONFERENCING?" 109
FIGURE 24: "HOW IMPORTANT ARE THE FOLLOWING FACTORS FOR EFFECTIVE AUDIO CONFERENCING IN EFL?" 110
FIGURE 25: "PLEASE RATE THE OVERALL EFFICIENCY OF USING AUDIO CONFERENCING." 111
FIGURE 26: "PLEASE RATE THE OVERALL EFFECTIVENESS OF USING AUDIO CONFERENCING." 112
FIGURE 27: "OVERALL YOUR SATISFACTION FROM USING AUDIO CONFERENCING IS:" 113

1. Introduction

The paper presents results of a field study investigating the user experience with audio conferencing in "Telenor Research and Innovations" (R&I). The current work deals with issues related to the use of information systems explored from a socio-technical perspective, including technology, information, communication, organization and people. The research methodology employed is ethnography, with data collection methods being interviews, observations and questionnaires. As analytical tool for systematisation and interpretation of the data, Activity theory (AT) was chosen because *"it can be used …to understand the way people act"* (Nardi, 1996).

More specifically the purpose of the study is to investigate the factors influencing the use of audio conferencing systems with focus on how the people are using the systems when communicating in English as a foreign language (EFL). The focus is on EFL, because *"understanding accented speech can be much more difficult than native speech, both because of the presence of an accent, and because grammar, pronunciation, and even word selection are much different than the listener expects"* (Rodman, 2006).

The findings identified in this work are not directly related to the issue how to improve technology in order to receive a higher QoE (quality-of-experience), as many other works investigating audio conferencing do [(Yankelovich, 2004), (Yankelovich, 2006), (Kellogg, 2006)]. The perspective through which the main issue is investigated here is a broader one: the idea is that technology, users of the technology, suppliers and facilitators of the technology are all interconnected and intertwined and this is reflected in the use of technology and the outcomes of this use. In other words the focus is not on the technology itself, but on the way it is used.

Audio conferencing is not only based on the technology, thereby the effort for solving only the technological perspective of the problem, is a one-sided effort. By applying

AT the aim is to investigate the problems accompanying audio conferencing as a complex activity. AT operates with diverse and at the same time systematic categories which offer the opportunity for a holistic[1] view of the analyzed issues. From this perspective Audio conferencing is a vast and versatile task for investigation. In addition, the current work is not aiming at solving all the problems related to audio conferencing, but to present a framework through which all the categories influencing the activity and the relationships between the categories can be seen. There is a huge amount of scientific materials discussing different issues related to audio conferencing but it is difficult to see the relationships among all the issues. The different disciplines are investigating the same problem from different angles and are often using different theoretical bases and notions. Therefore the concepts and categories offered by AT are used in order to organize this rich data in a matrix and diagrams consistent with the AT's theoretical framework.

1.1 Objectives and motivation

The objective of the study is to investigate whether there is room for improvement in the quality of the audio conferencing sessions. This could be done by illustrating how the audio conferencing is being adopted and integrated by people into their surrounding – a specific cultural setting. It is assumed that by improving the context (the conditions and circumstances) of the system, QoE will be improved. QoE could be enhanced through additional training of the participants in the meetings in using the equipment and by making them aware of behaviour guidelines that improve the quality of the meetings. Thereby the result of this research is aiming to provide requirements for support and arrangement of the cooperative work occurring in

14

[1] "**Holism** (from ὅλος *holos*, a Greek word meaning *all*, *entire*, *total*) is the idea that all the properties of a given system (physical, biological, chemical, social, economic, mental, linguistic, etc.) cannot be determined or explained by its component parts alone. Instead, the system as a whole determines in an important way how the parts behave.

The general principle of holism was concisely summarized by Aristotle in the *Metaphysics*: "The whole is more than the sum of its parts" (1045a10)." www.wikipedia.org

situations of multipoint audio conferencing in EFL. The aim is guidelines for appropriate meeting behaviour enhancing the QoE to be extracted.

As a **first step** the research aims at identifying the weaknesses of the communication practices (when using audio conferencing systems) in Telenor.

The second step is: The identified weaknesses of the investigated spheres could be improved through appropriate development methods. For example, improving the context (the conditions and circumstances) of the system will improve QoE, for instance: QoE could be enhanced trough additional training of the meeting participants in using the facilities and behaving in accordance with appropriate rules. Therefore the work has the objective to create a list of guidelines for effective user's behaviour before, during and after an audio conference which could improve QoE and enhance the outcome of the distributed meeting.

Audio conferencing is widely used in the companies nowadays and it has a tremendous impact on existing cooperative work practices. This fact imposes the need for deeply and thoroughly understanding this of type of work as a distinct category of cooperative work – articulation work. The articulation work is well studied in the field of CSCW (Computer supported cooperative work). With regard to CSCW it is appealing to investigate how this communication technology supports the way people cooperate, communicate and collaborate with each other and how they interact with technology. In this relation, Telenor as a large multinational company needs effective means of communication ensuring easy and effective flow of information between all the sub-units.

Specifically, for the purpose of this study, the factors influencing audio conferencing in EFL (ACEFL) need to be identified and the ways in which each factor influences the audio conferencing need to be understood. Some of the factors are very well described and quantified, in particular those related to technology. They have been studied through subjective testing of wide range of transmission parameters. All the parameters/ metrics, as for example latency, jitter, echo and packet loss, are included

in the so called E-model, that is actually a formula through which the R-factor value as specified in *TU-T Recommendation G.107* could be obtained.

"The E-model is based on a mathematical algorithm, with which the individual transmission parameters are transformed into different individual "impairment factors" that are assumed to be additive on a psychological scale. The algorithm of the E-model also takes into account the combination effects for those impairments in the connection which occur simultaneously, as well as some masking effects." (The E-model)

Through the R-Factor the QoE for VoIP (Voice over Internet Protocol) calls on a particular network could be assessed. The disadvantage of this method for assessing QoE is that it mostly works in laboratory conditions where the environment is strictly controlled. In the real world there are not two equal moments where all the parameters have the same value, although if the transmission is performed in evidently the same circumstances. It means that QoE could not be compared in two different real-world situations because the R-factor could not be obtained under comparable conditions. *"To the extent that impairments are present for which psychological additivity is not maintained, E-model predictions may be inaccurate"* (The E-model).

There are additional, complimentary factors which are not included in E-model and R-factor formula, as for example language, speed of speech, accent and dialect, which are enriching the complexity of ACEFL. Most of the factors investigated in this paper have subjective nature, such as language and dialect, but an attempt is made to quantify the factors with a subjective nature. This is done in particular by the rating scales in the conducted quantitative survey. In addition a qualitative description of the problems associated with those factors and their systematic analysis is done.

Audio conferencing falls in the category of the interactive telecommunication technologies. Even though video conferencing systems are more popular for investigation in research communities, this study explores audio conferencing

systems, because audio conferencing systems are still widely used in comparison with the video conferencing systems. They offer simplicity, easier and faster access and they are preferred for formal communication. There is quite a lot research done showing the advantages of video conferencing systems over the audio conferencing systems for particular communication situations. Some of the arguments for this are the following:

"Audio communication may reduce outcomes when used by people using a foreign language, compared with video communication",

"Video communication is preferred for non-urgent communication." (ETSI Guide EG 202 670)

This paper does not aim at proving those statements false or at stating that the audio conferencing systems in comparison to video conferencing systems are to be preferred, as many other investigations do. Both of them have advantages and disadvantages depending on the situations of use.

1.2 Research questions

At the very beginning of the study the research questions were not so clear, but they crystallised during the work process. This is because the study was not entirely oriented towards achieving a particular goal as for example revealing unknown truths - facts or relationships. The study was a spiral, cyclical process whereby every performed investigation aiming to study something very objective and oriented toward a specific goal (as for example the interview study, observation study and questionnaire study) new knowledge was acquired about the domain. The development of understandings was not only in a one predefined direction which had to lead (by plan) to the final goal, but the learning process came about by accumulating knowledge from the whole spectrum of ideas which support and accompany the main object of the investigation. This leads to a thorough

comprehension of the researched object, because in this way it is embraced together with its context (setting, environment).

The primary objective of the current study was to identify communication gaps among the distributed groups in audio-conferencing meetings where the attempt was to find methods through which to bridge the gaps between the distributed groups. Another objective was to improve the interactions across various group collaboration contexts in a real-world organization. These goals were identified prior to conducting the interviews. As everything in the nature undergo development and thereby changes, the objectives of the study also acquired new forms, which are directed towards the same goal, but they are more concrete and more deeply related to the context of the study.

These are the research questions:

1. What are the technical difficulties and what kind of approaches could be offered to overcome them?

   The technical issues are defined as those technical parameters that can be controlled only by the users, as for example calling to a conference, using the mute button, loudspeaker, microphone, testing the system etc. The network parameters are not perceived as a technical issue, because they cannot be controlled by users, but only by technical specialists. The issue was examined from this angle because the focus of the work is on user experience, which is also a main focus of investigation of CSCW and HCI.

2. What are the socio-cultural difficulties and what kind of approaches could be offered to overcome them?

   These are examples of socio-cultural issues: creating background noise, coming late for the meeting.

3. What are the linguistic difficulties and what kind of approaches could be offered to overcome them?

In order to answer these specific research questions, the following more general issues were also explored:

- How do the people use technology?
- How do the users perceive the technology?
- How do they learn to use the technology and what their interactions with the system lead to?

Notions like user experience and quality of user experience will be used in order to give answers to those questions.

2. Literature review

2.1 Audio conferencing

Audio conferencing (also called "teleconferencing") is the most commonly used form of conferencing since the 1930s when doing audio conference calls through the Bell System has been an option for businesses. Some of the reasons for that are that the phone is globally accepted and widespread today, easy to use and affordable. "*One in three UK companies currently use this type of 'green friendly' technology, according to recent research conducted by distributor Nimans"* (Conferencing).The following facts and numbers published in (Davis, 2005) are confirming the popularity and ubiquity of audio conferencing:

- Usage per year - over 20 billion minutes
- An annual unit volume growth rate - over 20%
- *"99.999 % reliability (often called the "five-nines of reliability"), the PSTN (public switched telephone network) is the gold-medal standard by which other communication services are measured."* (Davis, 2005)

Figure 1: Conference phone

2.1.1 Audio conference equipment

Audio conferencing systems turn the common telephone into a powerful meeting tool. Most of the common telephones allow users to use a "conference calling" feature. However, there are business telephone applications allowing audio

conferences of several calling parties with varying degrees of quality and diverse features. Each application is designed for a host of different environment – from a couple of users in small meeting rooms to dozens in large meeting rooms. Leading audio conference brands are Tandberg, Konftel, Cisco and Polycom.

2.2 Challenges in audio conferencing sessions

A number of theoretical and practical studies with a varied set of approaches to the situation have been undertaken to explore the challenges in audio conferencing (AC) with the aim to improve audio conferencing. The main object of most of the investigations in the telecommunication area are devoted to identifying and solving problems related to packet loss, jitter and delay effects. *"To date, there has been an implicit assumption in the networking community that many of these issues will be resolved through increased bandwidth"* (Watson & Sasse, 2000). Rodman (2006) argues that *"Of all the elements that affect the intelligibility of speech in telephony, bandwidth has been shown to be one of the most critical"* (Rodman, 2006). But Watson & Sasse (2000) continue their thoughts by arguing that *"If this is true, given the level of provision in the US and western Europe today, the quality of speech users experience in Internet conferences should be good. Yet, in a recent large-scale field trial, users reported speech quality problems in one out of three multimedia conferencing activities, where sufficient bandwidth was available"* (Watson & Sasse, 2000).

Watson & Sasse (2000) argue that users are rarely experiencing unsatisfactory audio quality due to packet loss and jitter. The experiments that the team conducted showed *"that in most cases, the impairment was caused by end-system hardware, equipment setup or user behavior"* (Watson & Sasse, 2000). Thereby the current investigation does not examine the purely technical aspect of audio conferencing, such as network parameters, bandwidth and protocol characteristics. The focus is on how the end user experiences these issues. Those aspects of investigation are the focus of interest for HCI and CSCW fields. In the present work the technical dimension of the

investigation was examined only on a level in which the user is interacting with technology – end-system hardware and equipment setup. The dimension where the end user does not interact directly with technology but his/her behavior introduces changes in the audio conferencing communication is looked as a socio-cultural dimension. The language issue, although it is included in the socio-cultural dimension is perceived as a different category, in respect to its importance for the study and the specific end-effect to communication process that it introduces.

Yankelovich (et al. 2004) reports that in their study *"the audio problems were the most frequent and had the largest negative impact on meeting effectiveness"* (Yankelovich, 2004). The research team identified that the **audio problems** that appear most frequently and are mostly correlated with meeting effectiveness are: some people could not be heard, poor audio quality, too much extraneous noise. They found out that the **behaviour problems** such as speakers not checking others for understanding, participants not following effective meeting behaviour, inadequate advance planning, and not facilitated meetings are highly correlated with meeting effectiveness. However, these kinds of problems did not seem to occur so frequently.

2.2.1 Technical challenges

Heim et al. (2001) categorizes the factors reducing the quality in real time person-to-person communication into three groups: network characteristics, codec characteristics and environment characteristics. For the audio conferencing sessions these factors are: The network characteristics for the packet switch network are bandwidth, packet size, packet loss, delay, delay jitter. The codec characteristics are the media protocols, delay and distortion. The environmental characteristics are acoustic and audio echo degradation.

Hestnes et al. (2003) identifies the important quality parameters which reduce the quality in the packet switch networks. For the input-output devices microphone and loudspeaker the authors describe the following:

"The microphone reduces the bandwidth, introduces distortion and mixes sound with the room's (or environment's) acoustic characteristics." (Hestnes, 2003)

"A loudspeaker introduces acoustic distortion in the mechanical parts, reduces the bandwidth and mixes the sound with the environment's (e.g. room's) acoustic characteristics." (Hestnes, 2003)

Noisy connection

The effect of packet loss is the most frequent technical cause for noisy audio connection. For solving this problem a quality enhancement mechanism concerning the terminal is proposed by Hestnes et al. (2003):

"For audio, as packet loss will result in noise it is common to decrease the volume of the output when data is lost; i.e. instead of playing the sound as normal with noise in the audio stream, the volume is decreased when one or more packets are lost. Therefore the user will experience a 'clipping' in the sound instead of arbitrary noise. This is considered less annoying for the user and tests have shown that even a relatively large amount of packet loss does not impact the communication." (Hestnes, 2003)

Bandwidth limitation

The upper frequency limit of the digital telephone network is defined by G.711 to be 3.3 kHz; while the lower limit is 220 Hz. In contrast the audible range of the human ear is approximately 20 Hz to 20 kHz. Most often the frequencies for speech occur in the "mid range" between 250 Hz and 8 kHz. Rodman (2006) argues that *"two-thirds of the frequencies in which the human ear is most sensitive and 80 percent of the frequencies in which speech occurs are beyond the capabilities of the public telephone network"*. This large gap between the human ear capabilities and the telephone performance affects the intelligibility of the speech carried by the telephone network. Speech intelligibility is discussed in section 2.2.3.

2.2.2 Socio-cultural communication challenges

"Social scientists have long been interested in small group collaboration, trying to answer questions such as: How can we assist groups to be more effective? Why do distributed groups and co-located groups perform differently? What impact does dominant behaviour have on group dynamics and performance? Group dynamics have been the focus of increasing interest as they are a key factor affecting the performance and satisfaction of the group." (Shaw, 1976)

Yankelovich et al. (2004, 2005, 2006) and her colleagues from Sun Microsystems have conducted a series of investigations of the challenges in audio conferencing sessions, with a further aim to improve audio conferencing. They have identified 10 top problems, where 5 of them were the following:

"Some people could not be heard ..... 33.9%..P=.004

Difficult to identify who was speaking..... 29.1%..P=.000

Poor audio quality ..... 23.8% .P=.001

Too much extraneous noise ..... 20.2% .P=.000

People did not follow effective mtg. behaviors .... 8.6%....P=.008

Percentages above indicate the frequency of a problem and the P values indicate the degree to which a problem is correlated with "meeting effectiveness." P=.000 is the most highly correlated." (Yankelovich, 2006)

The team has designed and developed an audio conferencing system called Meeting Central with the aim to solve some of the issues pursuing the audio conferencing sessions, by including information about participants, enhancing audio quality and the ability to spatially separate voices, providing text and speech based backchannels. (Yankelovich, 2004), (Yankelovich, 2005), (Yankelovich, 2006)

Another similar study was undertaken by Kellogg et al. (2006) from IBM Research Center with the aim *"to enhance the user experience and effectiveness of distributed meetings"* (Kellogg, 2006) and to lower the cost of audio conferencing. The IBM's approach is called Randezvous and it uses VoIP in order to lower the costs. The focus of the work is on two innovations which function as digital backchannels. The physical backchannels activities, for instance note-passing, whispering, presence (or attendance), are easy to grasp in face-to-face meetings in contrast to distributed meetings, where the backchannels might not be obvious.

Those examples show that the idea about improving audio conferencing is not new for exploration, but it is attractive, because there are a variety of different ways of achieving it.

Backchannels

Kellogg et al. (2006) describe the concept of linguistic backchannel in relation to audio conferencing. Linguistic backchannel is usually provided by *"utterances such as "uh-huh" or "yes, quite""* (Kellogg, 2006). Cogdill et al. (2001) analyse that the backchannel utterances fell into five categories: *"process oriented (analyzing or steering the group process or commenting on the experience of being in the online environment), content-oriented (private responses to mainchannel conversation), participation-enabling (helping users function better in the environment in which the interaction takes place), tangential (taking a thread started in the mainchannel offline), and independent (unrelated to the mainchannel conversation)."* (Kellogg, 2006)

Multitasking

"A recent survey revealed that 90% of audio conferencing participants multitask during their audio meetings instead of focusing on the topic at hand.[2] *"* (Davis, 2005)

[2] 2004 Russell Research survey of 385 audio conferencing users (www.russellresearch.com).

Verbal and Non-verbal cues

"...audio conferences do not allow participants to share non-verbal cues.... Mehrabian (1972) found that only 45% of the emotional meaning of a message is communicated by a person's spoken voice. The remaining 55% of a message is communicated through non-verbal cues including gestures, posture, and facial expressions.[3]" (Davis, 2005)

Selection of communication media combination

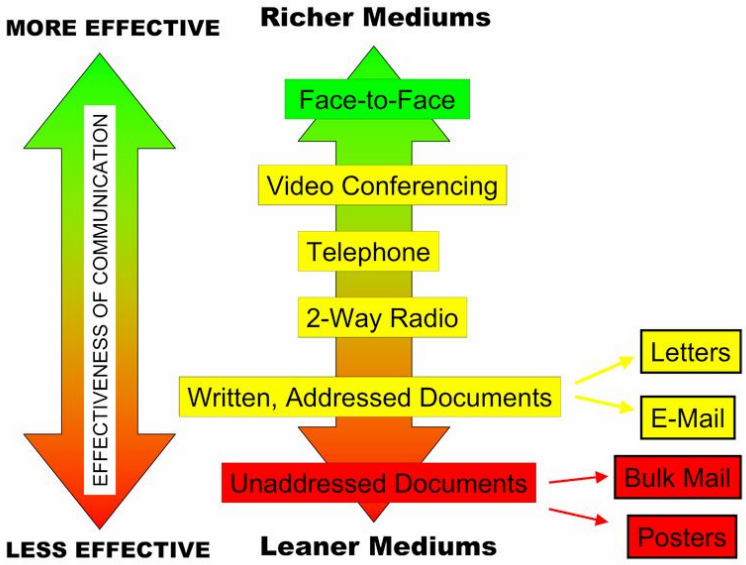


Figure 2: Explanatory diagram, Media richness theory; source: wikipedia.org

Heeren & Lewis (1997) have carried out a research centered around the factors contributing to the success of a distributed research community. He argues that the social factors playing an important role for the success of a distributed community are organizational and management structure and personal motivation, while on the technical level, which is the supporting bottom level, the success is influenced *"by*

[3] Mehrabian, A. (1972) Nonverbal Communication, Aldine-Atherton, Chicago, IL

the media selected for carrying out the activities in the community" (Heeren & Lewis, 1997, p.95). Heeren & Lewis (1997) have used AT in combination with Media-richness Theory as a framework to understand the context of work of distributed research communities and its developmental character. In addition, Heeren & Lewis (1997) found out that there are different needs for communication media during the different stages of the work process. The work processes are perceived as dynamic processes where the change and development is provoked by contradictions and frustration. Through the application of AT Heeren & Lewis (1997) identified three main levels of the work processes (or of the activity, according to the concepts of AT) – intentional level, functional level and operational level. *"The intentional level is oriented toward motives: needs, desires, or values… The functional level is oriented to specific, conscious goals in the context of the motives… The operational level is oriented to the practical conditions of actions"* (Heeren & Lewis, 1997, p.89).

Figure 3: The three main levels of the work processes, according to Heeren & Lewis (1997)

"Examples from the case studies show that using a rich medium such as face-to-face meetings is important in the startup phase of a project, so that shared understanding at the intentional level can be developed effectively. Results of higher-level activities will provide context for later lower-level activities which will thus involve less equivocality[4], and which will therefore require less rich media for effective and efficient collaboration." (Heeren & Lewis, 1997, p.95)

[4] Equivocality – "information may be interpreted in different ways" (Daft, 1987)

2.2.3 Language challanges in distributed communication

Herbsleb et al. (2000) have conducted a study in a geographically distributed organization. The results showed that when the communication was directed toward a native speaker of a different language and the face-to-face communication was not possible the persons preferred to use e-mail rather than the telephone. When the communication was directed toward a native speaker of the same language and the face-to-face communication was not possible the people agreed that the telephone is less ambiguous than electronic mail. These are the results of a qualitative study investigating the relationship between media preferences and native language presented in (Herbsleb, 2000).

On technical level Ren (et al. 2008) and his/her colleagues investigate *"the effects of different language on speech quality in voice over Internet protocol (VoIP) scenarios"* (Ren, 2008). They are motivated to do this investigation, because VOIP speech quality prediction or evaluation models are usually based on English language and *"the formula of these model are not really fit VOIP environment in our country"* (Ren, 2008). The team carry out an experiment *"to investigate how different languages (English, Chinese) effects on perceived voice quality on different VOIP system factors, which include different kind of delay distribution, codec and packet loss level"* (Ren, 2008). The authors suggest a new parameter related to language to be added to E-Model (ITU G.107). They argue that "*the enhanced E-Model achieves an improved measurement accuracy under different talker condition"* (Ren, 2008).

Speech Intelligibility

A study about the effect of bandwidth on speech intelligibility is conducted by Rodman (2006). Many of the issues in the aforementioned article although presented in the language chapter in the current work are tightly related to the so-called audio (or speech) quality issues by other authors (for example Yankelovich et al. (2004, 2006) and Watson & Sasse (2000) respectively). Thereby the article could be defined

as an answer to the question how to improve the audio (or speech) quality, which is a focus of interest by number of studies, including (Yankelovich, 2006).

Rodman (2006) describes that the consonants have a very critical role in the English language. Take an example with the words *pole* and *bole*: they have the same vowel, but the consonants are different; the different meaning of the words, when they are heard wrongly, can drastically change the meaning of a sentence. In addition the consonants occur more frequently in the speech. *"This critical role of consonants in speech presents a serious challenge for the telephone network. The reason for this is that the energy in consonant sounds is carried predominantly in the higher frequencies, often beyond the telephone's bandwidth entirely. While most of the average energy in English speech is in the vowels, which lie below 3 kHz, the most critical elements of speech, the consonants, lie above. The difference between "f" and "s," for example, is found entirely in the frequencies above 3 kHz; indeed, above the 3.3 kHz telephone bandwidth entirely"* (Rodman, 2006).

Other factors reducing speech intelligibility are reverberation, accent and the soft or whispered speech, according to Rodman (2006). The technological perspective of the language issue could be solved if the bandwidth increases to 7 kHz and beyond. The extended bandwidth will not only increase the speech intelligibility, but it can reduce fatigue and improve concentration, according to Rodman (2006). The author adds that also by *"extending telephone bandwidth below 300 Hz brings a significant increase in presence and realism"* (Rodman, 2006).

2.3 Relationship between QoS and QoE

Quality of service (QoS) and QoE are two important concepts in the current study. QoS could be defined as *"the collective effect of service performance which determine the degree of satisfaction of a user of the service"* (ITU-T E800). QoS implies when identifying problems from the service providers' domain, and includes the network measurements. Hestnes et al (2003) identifies that the QoS parameters

essential for audio conferencing in a packet switched network are delay and packet loss.

QoE implies when identifying problems from the customers' domain and measuring QoE is usually done by survey and subjective tests. Quality of Experience (QoE) may be defined as *"The overall acceptability of an application or service, as perceived subjectively by the end-user. Quality of Experience includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc). Overall acceptability may be influenced by user expectations and context"* (ITU-T, P.10/G.100).

Muhammad (et al 2006) describes the relationship between them as follows: *"the aim of the network and services should be to achieve the maximum user rating (QoE), while network quality (QoS) is the main building block for reaching that goal effectively"* (Muhammad, 2006). The following chart illustrates the relationship between QoS and QoE.

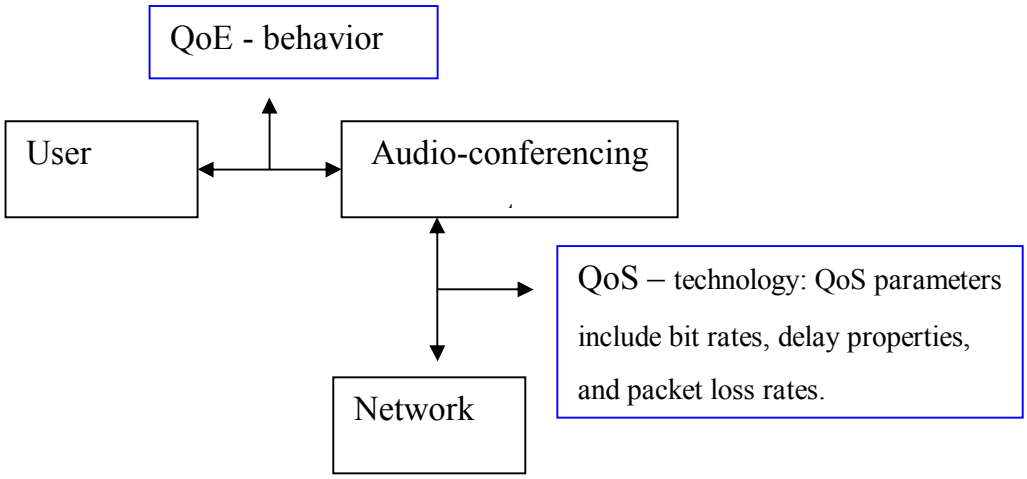


Figure 4: Relationship between QoS and QoE

2.4 Study of user experience

User eXperience (UX) is about how a person experiences using a system. *"User experience highlights the experiential, affective, meaningful and valuable aspects of Human-Computer Interaction (HCI) and product ownership, but it also covers a person's perceptions of the practical aspects such as utility, ease of use and efficiency of the system. User experience is subjective in nature, as it is about an individual's feelings and thoughts about the system. User experience is dynamic, because it changes over time as the circumstances change"* (User experience).

***"User experience (UX) evaluation** means investigating how a person feels about using a system (product, service, non-commercial item, or a combination of them). It is non-trivial to evaluate user experience and come up with solid results, since user experience is subjective, context-dependent and dynamic over time"* (Law, 2009).

There is not a specific approach for studying user experience, because there are several possible choices for each of the components of the study; it depends on what kind of system or product will be studied, where it will be studied – in a laboratory environment or natural settings; what is the purpose of the study – for example it could be early design of an artefact, or evaluation of a working prototype or refining and maintaining a product or an exploration of a new design concept.

In HCI there are three main approaches for evaluating user experience: usability testing, field study and analytical evaluation. In the current work the field study approach is applied. A characteristic of field studies is that they are conducted in natural settings with the objective of understanding what people do in their everyday life and how products mediate their work. In particular, the field studies can be used to: *"1) help identify opportunities for new technology; 2) establish the requirements for design; 3) facilitate the introduction of technology, or how to deploy existing technology in new context; and 4) evaluate technology"* (Rogers and Sharp, 2007, p. 591 – 592).

3. Theoretical framework

The fieldwork for this study was started without any pre-defined theoretical framework. The reason was that the original goal was to do an empirical laboratory test study with only a short ethnographic pre-study. The pre-study that was comprised of semi-structured interviews showed that the technological cannot be separated from the social. The interviewees were discussing not only the quality of the technology – the call, the audio conferencing system, the network; they were also discussing the social aspects of the communication – the culture and the language, about their and their colleagues' and partners' intentions, the social constructions – work groups etc. It became clear that the starting point of the investigation had a socio-technical nature, not purely technical and this reflected and changed the focus of the work. The initial way of seeing the audio conferencing was purely technical - ignoring the people, their work environment with its history and culture. Through the interviews the technology users were admitted to enter the "control room" where only the technicians are allowed to control the machines' parameters. Hereafter it was undertaken investigation in the same direction as the users' thoughts were leading. It was accompanied by a search for a theory which supports the aforementioned idea about the socio-technical nature of the work. There were several theories candidates for the analysis, which examine technology together with its socio-cultural context.

3.1 Identifying three main perspectives

After the pre-study three main perspectives through which the main issue could be examined were identified – technical, socio-cultural and linguistic. The linguistic perspective is also under the socio-cultural category, but with regard to the focus of the study which is audio-conferencing sessions especially in English as a foreign language, it could be seen as an additional category. These three categories are interwoven in the investigated field and any of them could be used as an analytical direction. There was an assumption that they form a sufficient union of viewpoints

which give the opportunity for a holistic understanding of the investigated domain. Aforementioned results from the interviews were accepted as a confirmation of the slightly implicit idea. Thereby these three directions are used as a main starting point of the investigation and after that in the analysis. These three perspectives are also used in the questionnaire, where the questions are separated in several subgroups including also technical, linguistic and socio-cultural groups of questions.

The next step of the work was to find a theory to be used as a lens through which to look at and analyze the investigated domain. The theory should be able to explore all the issues from two different and mutually complimentary (from holistic point of view) perspectives – technological and socio-cultural. Here the language perspective has a specific role: it could be included in the both - technological and in the socio-cultural level of analysis. Hereafter aroused the following questions: How the language issue is solved in technological level? How the language issue is solved in socio-cultural level?

On the socio-cultural level the language issue was examined by investigating the communication challenges that the users face when communicating with subjects with different socio-cultural background.

In order for the language issue to be examined on a technological level consideration was given to an experiment which could investigate how the available technology (the audio-conferencing systems) serves companies the needs of communication in different languages; is there any differences when the system serves for transmitting different languages, what are the optimal parameters of the network for transmitting non-native English speech. A similar study by Ren (et al. 2008) examining the Mandarin language in VoIP environment is presented in the *Literature review* chapter.

After considering the following arguments it was decided that such an endeavor is a task for an individual investigation and that it could not be covered enough thoroughly by the present work :

"Performing empirical studies normally demands substantial resources in terms of labour and equipment" (Heim, 2003).

"A team rather than a single person should in most cases carry out the study. A multi-disciplinary team is a great advantage, especially if it has both technical competence and competence within human factors or neighbouring disciplines" (Heim, 2003).

3.2 Searching for a socio-technical framework for the analysis

The socio-technical issue is covered to a different degree by a number of disciplines. In CSCW the socio-technical issue is the main focus of investigation. For the present work the following several theories used in the field of CSCW were examined:

- Activity theory (AT) was chosen as one of the candidates for theory, because *"Activity theory is most often used to describe activity in a socio-technical system"* (Bryant, 2005); some works examining AT and applied to the present paper are (Kaptelinin & Nardi, 2006), (Kuutti, 1991), (Bardram, 1998).

- Actor-network theory (ANT) is a sociological theory developed by Bruno Latour, Michel Callon and John Law and adopted by number of disciplines investigating technological domains. ANT is examined in many papers, some of which are (Aanestad, 2003), (Berg, 1997), (Hanseth & Monteiro, 1997).

- Strauss' Theory (ST) in (Schmidt, 1992), (Fitzpatrick, 1995), (Gasser, 1986). ST is not examined in great details here, considering the limited volume of the present paper.

- Distributed cognition (DCog).

Those theories are also widely used in HCI and Interaction design. There are other relevant approaches (includding theories, conceptual frameworks, descriptive methods, and variety of hybrid forms adopted from other scientific disciplines) used by CSCW and HCI that will not be examined here: conversation analysis in (Sacks, 1978), (Frohlich and Luff, 1989), (Sacks, 1992), (Katzenberg and McDermott, 1994), coordination theory in (Schmidt and Simone, 1996), (Carstensen and Nielsen, 2000), ethnomethodology in (Garfinkel, 1967), (Button, 1991), grounded theory in (Glaser and Strauss, 1967), (Strauss and Corbin, 1998), situated action in (Suchman, 1987), (Schiff, 1997).

All of the aforementioned theories were examined in order to select the most relevant of them with regard to the objectives of the study. The selected theory was to be used as a framework, through which the investigated issues from the domain will be selected, ordered, explained and analysed systematically. The first and most important criterion was that the theory specifically examines the socio-technical nature of the investigated phenomenon. This perspective benefits the research because it offers a rich and context oriented analysis of the selected issues where the infrastructure, which is usually left out of the technological achievements, is involved as well.

The following chapter briefly outlines the main characteristics of each of the selected theories, and at the end summarizes and defends the choice of the theory that will be used in the analysis.

3.3 Activity theory

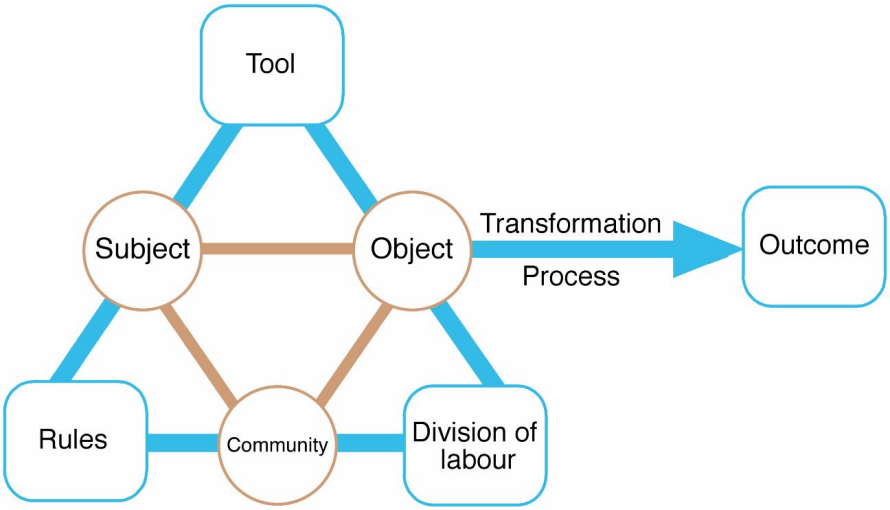


Figure 5: Engeström's (1999) activity system model

AT has been developed initially by Vygotski in the beginning of 20th century. It is part of Soviet Union psychological traditions and in its original form is a theory for cognition. AT has undergone changes in the 1990s with the works by Engeström (et al. 1999) and Nardi (et al. 1996), where the "individual" triangle of the activity system (with components subject, activity, and object) has been widened to a triangle including additional concepts – tools, rules, community, and division of labor. The *"later versions of the AT framework have become popular tools amongst researchers and some practitioners as a way of identifying needs for new tools and to analyze problems that are present in a work or other setting"* (Rogers and Sharp, 2007). First Bonnie Nardi and Kari Kuutt have applied AT to IS (information systems) and after that it has become widely used as a conceptual aid to research and design purposes and practices in the information and communication technology sciences, as for example in (Nardi & Kaptelinin, 2006), (Nardi, 2002), (Nardi & Redmiles, 2002) (Engeström & Nardi, 1998), (Nardi, 1996).

Consciousness

Another characteristic of AT is the incorporation of the notion consciousness: *"The object of activity theory is to understand the unity of consciousness and activity…. Activity theorists argue that consciousness is not a set of discrete disembodied cognitive acts (decision making, classification, remembering), and certainly it is not the brain; rather, consciousness is located in everyday practice: you are what you do"* (Nardi, 1996).

Hierarchical structure of activity

The hierarchical structure of activity is represented by 3 levels – activity, action and operation. The idea is that the activity, which is motivated by an object, can be broken down into goal-directed actions. Those actions are conscious. At the bottom level are the operations or routinized actions which become automatic and unconscious and require little attention. *"For instance, dialing a phone number can be a conscious action, but implementing this action by pressing phone buttons can be performed automatically"* (Kaptelinin & Nardi, 2006). The elements of the activity are not fixed, but they are changing dynamically as conditions change. Thereby the operations are also dependent on conditions.

Figure 6: Hierarchical structure of activity

3.3.2 Developing an analytical approach based on AT

Actually *"AT does not present a clear methodological prescription for the description or analysis of behavior as a set of procedures to be followed.[…]This means that identifying elements will be highly dependent on individual*

interpretation" (Rogers & Sharp; 2007). Thereby in the present work it is developed and used a specific approach for the analysis derived from the AS-triangle. In practice by using AS's triangle (Figure 5) AT can offer tools for displaying the 'bits' of data within 7 different categories – tool, subject, object, rules, community, division of work and outcome. After that these categories and the relationships between them could be explored in order to find out where tension between elements exists. For practical purposes it has been created here the matrix displayed in Table 1 which consists of the combination of each of two elements from the AS-triangle.

	Subject	Object	Tool	Rules	Community	Division of labor
Subject	**Subject**					
Object	1. Subject and Object	**Object**				
Tool	2. Subject and Tool	6. Object and Tool	**Tool**			
Rules	3. Subject and Rules	7. Object and Rules	10. Tool and Rules	**Rules**		
Community	4. Subject and Community	8. Object and Community	11. Tool and Community	13. Rules and Community	**Community**	
Division of labor	5. Subject and Division of labor	9. Object and Division of labor	12. Tool and Division of labor	14. Rules and Division of labor	15. Community and Division of labor	**Division of labor**

Table 1: Matrix representing the relationship between each two AS's elements

The Audio conferencing can be seen as an activity, with elements respective to the AT system latest model that includes:

Subject

The subject represents the activity performer. In the current case the subject of the activity is a person, a group or groups of persons who are involved in the activity – those who are participants in the audio conferencing. In the Scandinavian activity theory the notion of collective subject was introduced in the 1970s and 1980s. In the Leont'ev's model of AT the activity has only individuals, not groups as a subject, but he refers to "joint labour activity".

Object

Object (of the Subject) represents that *"on which the activity is performed"* (Rogers & Sharp; 2007). Most simply seen the object of any audio conference meeting is communication, exchange of information or coming to an agreement, but usually the object is different, according to the purpose (or objectives) of the meeting. It could be presentations, social meetings, discussions, monitoring and managing work processes, project meetings with clients, work or meetings with colleagues and others.

Tool

Tool is the artifact used by the Subject to accomplish a task; this is the conferencing system itself which has a mediating role between the Subject and the Object. The idea of mediation is perceived as a fundamental characteristic of human development in the change from a direct mode of acting on the world to one that is mediated by artefacts. Mediating artefacts shape the way human beings interact with reality. By using tools which mediate work activities the nature of work also changes. The knowledge of using a tool is transforming the human minds and this transformation could be seen as a historical process accumulated in the socio-cultural level. A good example is the use of telephone as a tool. The telephone influences the nature of communication: it changes not only the external behavior of the subjects

communicating through telephone, but also their mental functioning. In addition the knowledge of using a telephone is accumulated and transmitted as a social knowledge.

Rules

Rules are *"a set of agreed conventions and polices covering what it means to be a member of that community (set by lows, patterns, managers, boards, etc.)"* (Rogers & Sharp; 2007); this is the meeting protocol (etiquette) or meeting code represented by shared (or non-shared) conventions for acting and behaving during audio conference sessions.

Community

Community is everyone who shares the same objective as the Subject; this node of the chart represents the social context of the activity, e.g. all the people involved in the activity – directly and indirectly. Community in this study is the employees of Telenor who are using the audio conferencing systems, as well as all the clients, partners, and other who are communicating with the subjects in Telenor Fornebu.

Division of labor

Division of labor is *"the primary means of classifying the labor in a workplace, e.g. manager, engineer, receptionist"* (Rogers & Sharp; 2007). It is represented by the roles or different tasks that the participants are given in a particular meeting – for example there may be someone who is chosen to be the meeting leader / chairman, or it may be that different participants were given the task to report on something on interest for the meeting, another for example is giving a presentation.

How AT corresponds with the objectives of the study and the research questions

In order to apply AT to the objectives of the study it is appropriate to express the initial objectives in terms and concepts of AT. The starting point of the investigation is to examine audio conferencing as an activity in order to understand the individuals

and the social entities they compose in their natural work environment. In AT "activity" is a fundamental concepts and a unit of analysis. By giving a new shape to the objectives of the current research which fit to the notions in AT it could be argued that the purpose of the study is, by focusing on the analysis around the main concept of "activity", to identify the problems related to the use of audio conferencing systems at Telenor. Thereby AT can be applied to the process of revealing of problematic spheres among the practices accompanying the whole activity, through identifying the tensions between the components of the AS.

AT could be used as an analytical tool through which the tensions within or between the elements of the AT system could be identified. For example it could be looked not only at the issues directly related to the individual elements – Subject, Object, Tool, Rules, Community, Division of labor, but also at the relationships between them: Subject and Object, Subject and Tool, Subject and Rules, Subject and Community, Subject and Division of labor etc. This idea is represented by Table 1.

Another aim of the study is the practical contributions of the current work, namely the results to be used for design purposes, which corresponds to the sentence *"If CSCW is to be taken seriously, the basic approach of CSCW research should not be descriptive but constructive"* (Schmidt and Bannon, 1992). It imposes the application of such a theoretical framework which could fulfill these requirements. According to Bonnie Nardi who is well known for her work in AT and her main interests are activity theory, technology and communication *"Activity theory works well with design because activity theorists have always tested their theories in practical ways and believe that application is an outcome of theory, not a separate activity"* (Nardi, 2005). Thereby, the work aims not only at identifying problematic spheres, but also at creating a list of guidelines of actions it is required to perform in order to achieve the objectives of the activity and to improve the outcome of the activity.

3.4 Distributed cognition theory (DCog)

The two theoris based on the notion of cognition, AT and distributed cognition theory (DCog), are examined. AT has its roots in the Soviet Union traditions, whereas the DCog was born in Western Europe's traditions.

DCog was chosen as a candidate for a theoretical framework because it is usually used for analyzing collaborative work and it focuses on information propagation and transformation which corresponds to the topic under investigation in the current work.

The naming of the theoretical concepts in AT is powerful and very well depicted and it is useful for and advantages the description, according to Halverson (et al. 2002). A correct naming, i. e. using a particular notion in a theory is of great importance when it comes to manipulating data in process of analysis. Considering this AT has a great advantage in comparison to DCog where only *"few theoretical constructs are explicitly named in DCog. Those that are discussed, either in Cognition in the Wild (Hutchins 1995b) or elsewhere, are not presented in a way that gives them same the rhetorical force of naming as seen in AT"* (Halverson, 2002).

The main advantages that AT has in comparison to DCog are the differentiation between people and artifacts, clearly distinguished categories, and their representation over the Activity system diagram

Activity system diagram represents fundamentally dynamic interactions between the main components. This representation provides both descriptive and analytical power. In comparison with DCog framework where the process is central to the analysis, *"there is no iconic structure applied to each situation. Instead, it is built into the process of analysis, and may or may not be represented in the products of that analysis"* (Halverson, 2002).

Another disadvantage of DCog compared to AT is that in DCog the people, environment and artefacts are regarded as one cognitive system and the same theoretical language is used for both people and artifacts. Whereas in AT there are

clearly distinguished categories and a differentiation is made between artefacts and people.

3.5 Actor-network theory (ANT)

ANT was born within the field of social studies of science and technology. Actor-network theory describes heterogeneous actor networks consisting of human and non-human actors. It means that the theory doesn't make a difference between human and non-human actors. The act that is carried out and all the influencing factors considered together are linked and produce a network. In this actor network technical and non-technical actors negotiate with each other. Those negotiations aim at stabilizing, achieving a balance in the system. Negotiations lead to translations which transform the network continuously.

If ANT would be applied to the investigated domain, then the human actors of this network are the persons which are communicating through the audio conferencing systems, together with the non-human actors – the network and its parameters, the devices – audio-conferencing devices, microphones, the audio-conferencing rooms, together with the tables and the chairs etc. The main actor which could be chosen according to the aim of the research could be:

- An audio conferencing system in Telenor

Here the actor-network consists of and links together all the factors taking part in and influencing the acts which the audio conferencing system performs. Aanestad et al. (2003) describes how a non human actor, which is the camera, can form an actor-network. This method for analysis manifested in the paper could be easily applied to similar situations, as well as for the present case by examining the audio conferencing system as an actor. The weakness of this method is that it could serve only for description of an actor-network system, by adoption of the concepts presented in the ANT. ANT does not present a systematic framework for analysis. The purpose of the

present research is to identify the consistent factors in the network and to depict the relationships between the elements.

A noticeable difference between AT and ANT is the incorporation of the notion consciousness in AT. According to Vygotsky 'consciousness' unifies attention, intention, memory, reasoning, and speech. *"Activity theory, with its emphasis on the importance of motive and consciousness - which belongs only to humans - sees people and things as fundamentally different. People are not reduced to 'nodes' or 'agents' in a system; 'information processing' is not seen as something to be modelled in the same way for people and machines"* (Nardi, 1996). In ANT the people are reduced to nodes or agents in the system, and 'information processing' is seen as something to be modeled in the same way for people and machines.

ANT doesn't make a difference between human and non-human factors in the unit of analysis. AT has a different conceptualization in the unit of analysis – the system is divided into 7 units (subject, object, tool, rules, division of labor, community and outcome). Another difference is that AT makes a systematical differentiation between the elements in the system, and orders them according to their relationship to the main subject of activity. Whereas in ANT all the actors are equivalent and they could change their role in the process of negotiation and translation of the system. In ANT the whole network and its actors could be totally transformed, while in AT the focus is on the main activity together with the system constituting of predefined elements which support the activity.

3.6 Strauss' theory

Action / interaction and *social worlds* - *"the fundamental building block(s) of collective action"* (Clarke, 1991, p. 131) are main notions in Strauss' theory (1993). A social world is an interactive unit that arises when a number of individuals strive to act in some collective way, often requiring the coordination of separate perspectives and the sharing of resources. It has *"at least one primary activity (along with related*

activities), ¼ sites where activities occur ¼ [and] technology (inherited or innovative means of carrying out the social world's activities)" (Strauss, 1978).

3.7 Conclusion

What value does each theory, or approach, provide

It should be underlined that not one of the theories is universal, there is not such a theory through which all the things could be seen, identified, understood, and explained because the theories are just frameworks through which the world, and respectively the world of a particular problem area, is described:

"by shifting theoretical perspective the world under investigation also changes shape" (O'Brien, 1993:10-11).

But the aim to find a theory which can provide relevant insights to the selected problem space is achievable. It is of a vital significance that objectives of the study and research questions are aligned with the theory applied for the investigation in a way that the theory highlights relevant issues and shapes the objectives. *"The value of any theory is not 'whether the theory or framework provides an objective representation of reality"* (Bardram, 1998). *"The benefits of the theory may be seen both in the structuring of the empirical data, and even more so through the leverage it allows in the interpretation of the case"* (Aanestad, 2003).

Through the application of ANT one can not achieve a good analytical part (where for example hidden sides and patterns are revealed and synthesized), because the level to which the theory reaches is only to generalize and synthesize (by applying the theory to the case), which is a process opposite to the analysis. Any investigation consists of a descriptive part and an analytical part, where the descriptive part aims to depict the problem area, while the analytical part is the cumulative point of the investigation which presents unrevealed to the current moment characteristics or relationships of the investigated issue/s. AT can support both – description and analysis. It supports

the description because an activity system which is represented graphically by a triangle consists of all-embracing and important nodes, important with respect to an activity. Thereby AT is very useful and has advantages in comparison to the ANT, DCog and ST that were described above in the current chapter.

4. Description of the field setting

In this chapter the field setting and the conferencing systems that are under investigation will be described. The chapter starts with a description of Telenor and its research department located in Fornebu.

4.1 Telenor

Telenor is a large telecommunication company with headquarters in Fornebu, close to Oslo, Norway. *"Today, Telenor Group is mostly an international wireless carrier with operations in Scandinavia, Eastern Europe and Asia, working predominantly under the* ***Telenor*** *brand"* (Telenor). Figure 7 shows the Telenor world locations for 2009 year. Telenor holds controlling interests in mobile operations in Norway, Denmark, Sweden, Serbia, Ukraine, Hungary, Montenegro, Thailand, Malaysia, Bangladesh, India and Pakistan. The fact that approximately half of the employees work outside Norway imposes the need for effective means for communication among all the subunits.

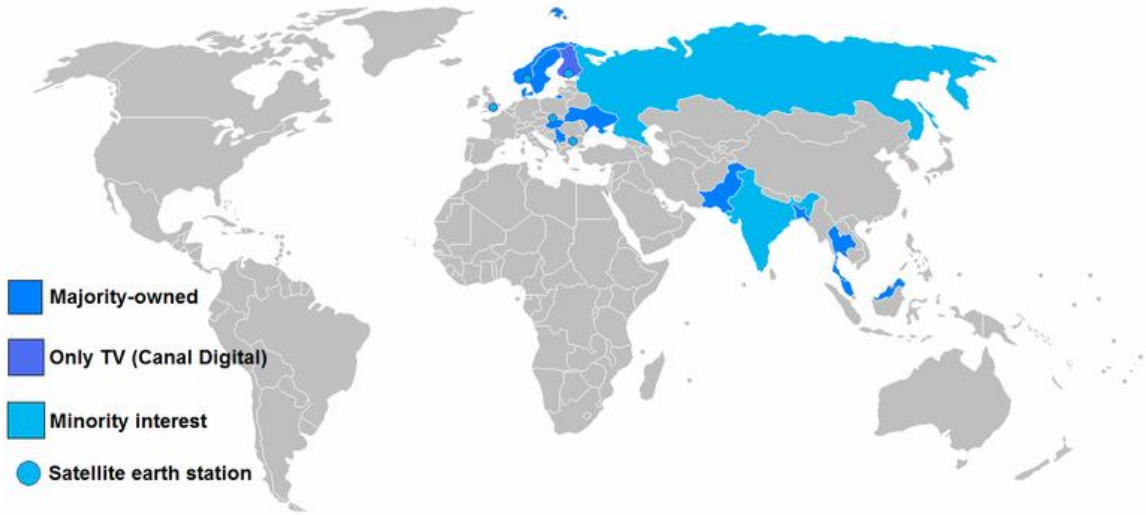


Figure 7: Telenor world locations 2009, source wikipedia.org

4.2 Audio conferencing systems in Telenor

In Telenor R&I the audio conferencing sessions are usually conducted from the facilitated meeting rooms. Conventionally all the telephone conversations are supposed to be conducted there. There is another internal VoIP-based conferencing system called "Communicator" which is supposed to be used both for data conferencing, video conferencing and audio conferencing. But when a single participant or a small number of participants from this department are communicating with another distributed group then they feel free to choose which system to use. For example some of the employees prefer to use Skype or NetMeeting (Microsoft) or a mobile phone to connect to their distributed colleagues or partners. For organised meetings where a big number of employees from R&I are participating the facilitated audio conferencing rooms are usually used. The videoconferencing facilitated rooms are rarely used in the department, because audio conferencing is preferred to videoconferencing, mostly because of the simplicity of audio conferencing equipment in comparison to videoconferencing equipment.

4.2.1 Audio conferencing systems using PSTN network

The audio conferencing meeting rooms are facilitated with audio conference equipment using the PSTN network.

The public switched telephone network (PSTN)

*"The **public switched telephone network (PSTN**) is the network of the world's public circuit-switched telephone networks"* (Public switched telephone network). The PSTN is almost entirely digital and covers 99% of the world. There are still analog telephone systems in a few countries.

Audio conference equipment

Figure 8: Conference phone by Konftel in one of the audio conferencing rooms in Telenor R&I

The audio conference equipment used in R&I department are professional business applications produced by Tandberg or Konftel. Through those phones a calling party from Telenor R&I could call several other calling parties and include them in a conversation. Another solution for initiating an audio conference session is all the distributed parties to call a special "bridge" number at the same time which could connect them to a conference bridge by linking all the calling telephone lines.

Conference bridges

The conference bridge is a public service offered by Telenor. The employees in R&I are using the audio conferencing equipment in the meeting rooms as an ordinary telephone in order to dial the conference bridge. Those telephones use circuit switching that routes calls through a series of interconnected switches until it gets the audio conference bridge. One of the characteristics of analog circuit switching is that all the connections stay open for as long as the session continues, though this can tie a network up significantly and results in the poor sound quality most people are familiar to in AC. In addition, since only one person normally talks at a time, a lot of excess capacity on the network is lost with this tool.

4.2.2 Internet applications

The Internet has replaced the ways people communicate with each other. The audio conferencing phone and audio conference bridge are replaced by cheap hardware running Voice over IP software.

Skype

Skype is a world-popular VoIP-software used as well in Telenor for conducting audio conferencing sessions. The traffic generated by Skype has the following characteristics: the bandwidth requirement is low - 69 Kbps and the packets are small - on average 236 bytes using TCP. To avoid that the user experience suffers, ITU-T has specified that the latency (one way) should not exceed 150-200 ms, and it must not exceed 400 ms (ITU-T Rec. G.114, 2003).

Communicator

The tool *Communicator* is an IP-based application specifically designed for sharing data when the participants are separated geographically. Through it one can share its desktop with the other and he can have video, audio and desktop simultaneously.

Microsoft Live meeting

Normally, for internal meetings with colleagues from Telenor Norway the Microsoft conferencing solution Live Meeting 2007 is used. It has audio and video functions and such for sharing multimedia content. One can use both, computer and telephone conference service to connect to a conference meeting held on Microsoft Live Meeting.

IP telephones

VoIP telephone systems in Telenor use one of the G.7xx formats recommended by ITU-T[5]. G.711 and G.729 have bandwidth requirements of 64 and 8 Kbps respectively, and the packet size is determined by the packet transmission cycle. This is typically a few tens of ms, resulting in packets that are between 80 and 320 bytes for G.711 (Hassan & Alekseevich, 2006). Other network parameters characterizing the networks at Telenor were not available during the current fieldwork, because of the security conditions imposed over the external individuals (see 5.5).

[5] International Telecommunication Union - http://www.itu.int/ITU-T/

5. Methodology

The research methodology used in the current work is ethnography. The data collection (research) methods are interview, observation and questionnaire. This approach of using various different research methods is called *methodological triangulation*[6]. By applying triangulation the weaknesses of one method can be avoided by using another method. For instance, in the interviews the interviewer could not be sure that a respondent is telling the truth until he checks the results by conducting observation or a survey. A data collected through only one research method is usually not as reliable as the data collected through combination of two or more research methods. Triangulation increases the control over the accuracy of the data.

The ethnographic field study is complemented by a quantitative survey in order to confirm the results by the qualitative study (interview study and observation study) and to substantiate the results by adding a numerical value to them.

5.1 Ethical issues

In ethnographic studies the ethical issues are of great importance, because the researcher as an external person is getting access to internal (and very often personal) information. It has been followed the Norwegian regulations on collecting and managing personal information entitled: Personopplysningsloven (Personopplysningsloven, 2000). In addition the information revealed in the research paper is consistent with the norms and regulations of the organization (Telenor) where the research is conducted. In the beginning of the field work a document called "The promise of secrecy" provided by Telenor has been signed. It assures that all the information that is accessed has to be used only for research purposes. Also all the

[6] Ricardson & St. Pierre 2005:963 presents a critical account on Triangulation

interviewees have been asked for consent to use the shared information for the paper and they have been informed about the conditions with which the data will be used, as defined in Section 19 in Personopplysningsloven. The interviewees have been assured that their anonymity will be kept and that the materials are collected only in relation to the topic of the thesis, in accordance with Section 11 in Personopplysningsloven.

5.2 Ethnography

Ethnographical research originates from the field of anthropology and sociology. The aim is to explore peoples' everyday practices and what is going on in a particular setting - organizations, subcultures or tribes. Ethnographic research is characterized as an in-depth and extended study, where the researcher immerses himself in the field in order to create a thick description of the investigated domain. Ethnographic study can give researchers an in-depth knowledge of the work practices. The focus of ethnography, as a sociological approach is to develop understanding of human behavior, which is done in practice by extensive contact with the people in the investigated setting. As a result from the field work an ethnographer develops descriptive understanding of the investigated behavior, which s/he represents in a form of thick descriptive narratives.

Another characteristic of the ethnographic study is that here pre-formulated study design is avoided, but epistemological discipline and systematic method are pursued. In addition the things that are taken for granted should be problematized, it means that insider's view is not necessarily the best and the task is not to replicate the insiders' perspectives. Also, the people's view should be turned into results, because it is not a result, but it is data which has to be understood, interpreted and analysed. (Forsythe, 1999)

Ethnographic data collection methods are observation, interview, video and audio recording, survey. The data collection methods applied in the current study are

observation, interview and survey questionnaire. While describing about the six misconceptions Forsythe (1999) also mentions that it is not enough just to ask people in order to reveal what they think. She argues that the predictive value of verbal representations and the generality of short-term observations are questionable and recommends complementary extended observations to be conducted. In the current field work there was not such an opportunity for conducting multiple continues observations, because of the specificity of the setting which is explained in the chapter *Limitations*. It was not allowed the attendance of the audio conferencing sessions in Telenor by external persons, because the shared information in those meetings is assumed to be secret and only for internal use. The results from the interviews could not be confirmed by observations (because of the aforementioned reason) and the only chance for collecting thicker and more precise data was through a survey. Thereby the ethnographic field study was combined with a quantitative survey.

The question about whether the ethnographic data collection methods include quantitative questionnaires is debatable, because it is assumed that ethnographic approach produces mainly qualitative accounts. According to other authors *"Questionnaires can be used to aid the discovery of local beliefs and perceptions and in the case of longitudinal research, where there is continuous long-term study of an area or site, they can act as valid instrument for measuring changes in the individuals or groups studied"* (Ethnography).

5.3 Ethnography in computer sciences

"The ethnographic approach, with its emphasis on "natives' point-of-view," holism, and natural settings, provides a unique perspective to bring to bear on understanding users' work activities" (Blomberg, 1993). The understanding of users' practices which the ethnographic methodology is offering is useful for and applicable to the design of new technologies. When the ethnographic approach is applied to a technological domain there is a shift of the priorities: the ethnographic approach is

applied for the purpose of understanding human behaviour in relation to technology – understanding the users' needs and their abilities of using technologies.

5.4 Relationship between research methodology and theoretical framework

There is a tight relationship between AT and ethnographic methodology. First and foremost because both - the ethnographic approach and AT provide an opportunity for holistic interpretation of the phenomenon under investigation:

"A typical ethnography attempts to be holistic" (Ember, 2006).

"Activity theory provides a holistic framework to investigate relationships among the elements" (Koszalka, 2004).

"Activity Theory is a tool, borrowed from the field of psychology and development work research, which facilitates a more holistic view to be adopted by information systems practitioners and developers" (de Freitas and Byrne, 2006).

Second, because the combination of AT and ethnography has typically been used for investigations in CSCW, HCI and Interaction design domains:

"The application of AT has been limited in domains such as medical practice (Engeström 1993, 1995), juridical praxis (Engeström 1996a), programming (Holland et al. 1996), software design (Gould et al, 2000; Turner & Turner, 2001) etc. In such domains it has offered rich accounts mainly of ethnographic character" (Nathanael, 2002).

And third, because the combination of AT and ethnography provides useful insights for design, although not so often applicable, because of the following difficulties: *"i) to provide concrete recommendations at the level needed by designers, ii) to represent the causality that is manifested by the technical components of socio-technical systems"* (Nathanael, 2002).

5.5 Limitations

The field study at Telenor was accompanied by many limitations, which undoubtedly influenced the quality of the investigation. The primary source of limitation came from the security conditions imposed over outsiders presenting in Telenor. Technical information about the audio conferencing systems and services, about the networks and the networks' parameters was not available. Also an observation study could not be conducted.

How the limitations affected the study

The observation is the main data collection method in the ethnographic investigations, because the researcher herself is the main research instrument. Thereby the primary aim of the fieldwork was to conduct participant observations of the work processes and to investigate how the work is accomplished with the mediation of audio conference. Through the observation one could be a witness and can get rich data directly from the source in order to understand and describe the context of a social and cultural phenomena and situations. Judging on the basis of the researcher's own experience is important for the quality of the future ethnographical analysis. The different nuances and shades of the investigated objects could be identified and understood truly if they are seen and experienced directly from their source. But what have been acquainted from the work were only reflections from the respondents in Telenor - the accounts form the users through the interviews and the survey. The observations that have been conducted were out of the investigated domain but from a similar setting and it was considered as a valuable data because without it the perception of the investigated processes would have been poorer. As it is in any ethnographic study, the researcher herself takes decisions about which data to use and how to use the data, and thereby it was considered that the data from the observations can support the analysis.

The reason why the data from the interview and questionnaire study was not considered as primary data (as they would be if it was conducted an observation in

Telenor), and thereby not valuable for an ethnographic study, is because it is very often that the people are not objective. For example research has proven that interviewees are tending to describe phenomenon according to their own idealistic view (of the things) and the interviewees do not realize this tendency.

"The distinction between what people say and what they do is related to the distinction between ideal and manifest behaviour. Ideal behaviour is what every "good" member of the community should do. Sometimes asking people about their behaviour will yield responses closer to the ideal than the manifest. People may distort, either knowingly or unknowingly, accounts of their own behaviour, often simply providing an approximation constructed either for the questioner's benefit or to match cultural expectations" (Blomberg, 1993: 130).

6. Interview study

This chapter consists of the following parts: aim, participants, method, results and discussion. In the first part the aim of the interview study is explained. The second part describes the participants of the interview study. The third part examines the interview study as a methodological approach. In the fourth part the results of the interview study are discussed and summarised. The fifth part draws a conclusion.

6.1 Aim

The interview study was the first stage of the field work. The aim of the interview study was to explore how the users are seeing the audio conferencing systems and the issues related to them, to understand what kind of categories the interviewees are using to describe the audio conferencing. This data is important for the purposes of the further investigation.

The investigation aims at finding out what do the people think is the primary problem - the technical quality of the connection, whether the system is user friendly, efficient and effective or whether the meeting participants' behavior is adequate.

6.2 Participants

The participants in the interview study were randomly chosen employees in Telenor R&I. Telenor employees who were willing to participate in the study were met in advance in their work places and asked for consent and for an appointment for the interview. As a result of these prearrangements 12 informants[7] were interviewed.

[7] "Ethnographer's traditional term of respondent" (Blomberg, 1993: 136)

6.3 Method

Three series of open-ended, semi-structured interviews were conducted in Telenor research department in the end of April and in the beginning of May 2009. The interviews had some characteristics of contextual ones because they were conducted in the audio conferencing meeting rooms in the research department in Telenor, which are also used as face-to-face meeting rooms. Only one researcher was interviewed in another setting which was more convenient for him. The contextual interviews can provide richer contextual information than the usual interviews. In the current situation the interviewees were asked to show which facilities they are using for audio conferencing and how.

As usual *"the ethnographic interview is not bound by explicit rules. To the contrary, a great deal of the latitude is given to the interviewer to exploit the particulars of any given interview situation"* (Blomberg, 1993: 136). The only rule that was followed was the interviewees not be interrupted and to be left to lead the conversation in a direction appropriate to their view, because *"the objective of the interview is not to produce the most accurate or complete understanding of the topic, but to gain a better understanding of what the respondent knows and thinks about the topic"* (Blomberg, 1993: 137). It was not known in advance how the users perceive and how they talk about the system. They were consciously not asked direct questions, which lead to a certain type of answers. The questions in the conversation were very carefully put so that the users were encouraged to use their own words and categories in the description of the issues related to the systems. The interviewees were given the opportunity to shape the content and the character of the conversation.

There is little volume of discussions in the qualitative research literature on the topic for audio recordings of interviews. Modaff et al. (2000) provides tools and techniques for audio recordings of interviews. The factors that have been considered with regard to the interviews are the following: iPod has been chosen because it is an easy to use device and provides a good audio quality of the recordings; the audio recordings have to be easily transferred to a computer - the audio format should be compatible with

and easily reproduced on a computer; an external microphone attached through a mic-in jacks has been used; control over the recording process was considered together with copy protection.

Unexpected recharging of the battery had to be done twice during two different interview sessions. It was caused by reasons out of the researcher's control, and after the battery was recharged the interviewing continued.

Transcription of the interviews

The audio quality of the recordings made with IPod is good and this advantage decreases the time for transcription. The transcription of the interviews is time consuming process and one hour of recorded interview takes 4-6 hours for transcribing. 12 interviews have been taken for about 30 minutes each, which makes approximately 6 hours recorded audio. These 6 hours audio are equal to not less a 36 hours time spend for transcribing. After having the opportunity to experience this time-consuming process and after the transcription of the third interview the decision was made to search for software which can minimize the transcription time and support the finalization of the task. Two types of software have been found, installed and tested (whether they can work properly) - 'Wave to Text'[8] and 'Transcriber'[9]. However, after trying them both out, they were not found to be useful. Further efforts to find better software were fruitless. What was found after the long search was the article explaining this phenomenon as follows:

"You can't do this at any price: even humans still have problems transcribing interviews accurately. After more than 25 years development, Nuance's Dragon NaturallySpeaking software can turn speech into text, and it can handle digital files. But it only works well if you have a clean file (without much hiss or background noise), and if the user speaks clearly and carefully" (Schofield, 2008).

[8] http://www.research-lab.com/vexp007read.htm

[9] http://trans.sourceforge.net/en/presentation.php

"Why can't I use voice recognition software to create a text document? While it would be great to avoid the time-intensive task of transcribing interviews, voice recognition software does not yet work well enough. The software has to be "trained" for each voice, and speakers must speak slowly and enunciate clearly" (Forester, 2005).

6.4 Results and discussion

In this chapter the emphasis is on searching for and interpreting themes, patterns, features and relationships from the data following the main research question - what kind of challenges the users of the audio conferencing system encounter. Parallel with it the themes shall be refined and collated.

The status of the data is such that first it has been ordered by the interviewee's construction and afterwards by the construction of the current analysis. It means that the data has been gone through two lenses – the interviewee's and researcher's lenses, until it is represented here in the current format.

The results from the first interviews have surprisingly showed that the three perspectives – technical, socio-cultural and linguistic were clearly distinguished by the users. In confirmation of the initial idea that these three categories form a sufficient union of viewpoints offering a holistic understanding of the setting, some of the conversations with the researchers went according to the same distinguished categories, although it was not supported in all of the cases. It is supposed that the researchers' and the author's viewpoint of seeing and distinguishing these three perspectives is formed by similar education background (the theoretical knowledge gained through a course of education in the fields of social science, information systems, human-computer interaction and CSCW is similar) or just from their practice (while reading and writing research articles). There is an assumption that if the same investigation has been made in a different department in the same company, but where the information systems and issues related to them are very much the same as

in R&I, the users would have talked differently (using different words and concepts) about the same issues. In the current situation the concepts used by the researchers are very similar to the concepts through which the problems in this paper are described. Despite this fact which could be accepted as an encouraging for the investigation, the practice has shown that information systems implementation fails without any differences in respect to who the users are. There are examples describing failure in information system implementation in research departments where the users are very well acquainted with all the issues related to the particular information system.

6.4.1 Technical issues

A common meaning shared by the interviewees was that the combination of all the technical problems makes it more difficult in understanding accented speech and EFL. The audio quality was mentioned as the most often problem during audio conferencing. However, the lack of knowledge for setting up an audio conference was found to be the critical point diminishing the meetings effectiveness.

Setting up an audio conference

Another issue is setting up audio conference. The lack of knowledge of using the devices reflects to late start of AC and also decreases QoE. The following passage from an interview illustrates this issue:

"We were a group of people and had to have a teleconference meeting. But no one knew how to set up the teleconference. We spent about 20 minutes until we connected to the meeting. But then the purpose of our participation in the meeting was not fulfilled. And the quality of the meeting dropped, because we were not able to discus the issues that the other participants were discussing."

It was discovered that some of the people in the department did not know how to set up an audio conference, because there is not information in form of instruction about how to use the devices. In addition the devices in the different rooms are not from the

same type. Some of the researchers were used to using a particular device, but they did not know how to use the other. Thereby they shared that the devices should be consistent in all the rooms.

Complexity of devices versus effectiveness of devices

What makes an impression is that each of the users has their own preferences for what the system should consist of and how it should be used. For example, a technical solution for recording the meetings to be integrated in the audio conferencing in Telenor was suggested by a respondent. It will support the employees' work by helping them to recall situations and listen to the discussions from the past.

Integrating all the suggested tools in a one audio conferencing system is proved to be undesirable solution from users with regard to the complexity that such systems create and the problems they would introduce in the process of use (Yankelovich, 2004). Instead of increasing the complexity of the tools while aiming at improving the audio conferencing the development would rather be directed to maximizing the meeting effectiveness, because *"people want to focus on the content of their meeting, not on the meeting tool"* (Yankelovich, 2004).

6.4.2 Socio-cultural communication issues

Background noise

The background noise is rated as one of the worst factors decreasing the quality of the meeting. An interviewee from R&I shared his experience during an audio conferencing session including 3 locations and with about 10 people:

"The noise was too much and with the overlapping it was almost impossible. In the some of the places there was lots of background noise like, you can hear cars and others like that. In one of the other locations the person was working from home so there was a lot of noise made by the kids and he had to apologize all the time."

A very common problem during audio conferencing meetings is that the people are not aware that they are creating noise when they are moving papers or typing on the computer, and are close to the phone (or the microphone). This noise is perceived as very annoying for all the distributed participants with exception for the co-located attendees.

The making of background noise sometimes could be considered as, but often it could not be considered as a behavioral issue, because of the fact that the participants in the meetings are not able to control their environment or to find a better one. Not all of the distributed attendees have isolated rooms like those in Telenor R&I. However, all of them could control the environment to a certain extend. For example they can close the windows, doors, etc to isolate the external background noise enter in the communication channel. They also can be made aware that the noise that they are creating by their movements is directly picked up by the microphone(s) and volume turned up. In the case when the attendees are participating in the meetings through their mobile phones because they are somewhere outside the office it is not possible to isolate themselves from the noisy environment. Then QoE (the quality of experience) suffers.

From technological point of view the problem with the background noise is solved through noise cancellation mechanisms integrated in some of the table top conference telephones (and Figure 8: Conference phone by Konftel in one of the audio conferencing rooms in Telenor R&I)

Whispering

A researcher shared: *"One of the problems in audio conferencing is that you do not hear yourself, you can not hear the voice that you are making. It is the other ones hearing it. For example the whispering; whispering is one of the worst things that you can do. Because if you are several people in the room, you and I should whisper together in order to other ones not hear, but the ones on the phone line, they hear*

because it is going straight into the microphone. That is one of the worst things that you can do."

Selection of communication media

One of the respondents argued that *"teleconferencing is not a substitute for face-to-face meetings, but it is an additional medium of communication"*. He said that:

"When it is ok to use audio conference it is when you know the people on the other end. If you do not know them at all it is a bit awkward. I would prefer to meet the people at least once."

Another researcher described what the preferences for communication media in the department are:

"It is typical for us to use teleconferencing and not videoconferencing or data-conferencing. And we are usually using this equipment in the isolated rooms. Sometimes we can use the tool called Communicator, specifically designed for sharing data when we are separated geographically, but I do not think that this tool work that well. I have encountered problems using it; it is not so easy to use this application. The sound face, lots of time the connections are dropped in the middle of the conference. And they continue to drop and drop and drop. Then we decide to end up with an ordinary telephone or teleconference and to relay on data sharing only when it is really needed. I do not like to use the Communicator, because I do not find it reliable: the video is dropping and the voice is dropping. That goes through data network. I don't know whether the problem is in the application or in the network."

Regarding the preferences of using the audio conference equipment in the conference rooms, a researcher summarizes that it is dictated by the fact that *"when you use those devices most of the time it is ok"*.

The fact that the mobile phones tend to pick static is also experienced by some of the researchers:

"In the mobile device you have more noise than in the conference device."

Despite this fact a researcher shared that he prefers to use a cell phone to connect to audio conference meetings because with it he could be 'mobile' and could for example go to the coffee machine to get a cup of coffee.

Audio conference protocol / etiquette

An interviewee shared that the people that are using audio conferencing as a communication media often have a proper behavior during the meetings. They have mastered the fine points necessary for an effective flow of the audio conferencing meeting. *"Sometimes it is difficult to recognize the different voices when the people are speaking. But some of the people who are attending often teleconference meetings are used to use a teleconference etiquette and before they start talking they are saying their name, for example "Hallo, this is John", or "John is speaking"."*

This quotation and the following one illustrates that the participants in audio conferences are learning the meeting protocol from the practice and the appropriate behavior is becoming a skill.

"In the previous company that I was working for, it was not so big and was not international and I did not have to talk so often to someone else, there was immediate group of work. But when it happened teleconferences were "funny" events: people did not know what it is gone be. Because we were mostly used to talk on telephone with a loudspeaker, rather than having a teleconference equipment. Before teleconferences were more difficult, but now the people are used with the teleconference experience in the everyday work process and also in their private live."

Importance of group management structures

Heeren & Lewis (1997, p.95) argue that *"human moderation and leadership of group interactions are critical concerns which must be understood and expressed explicitly in order to maintain group cohesion"*.

One of the researchers, an experienced meeting leader described the most important tasks that a meeting leader has to manage with and what has to be taken into consideration when taking on this role. First and foremost the meeting leader has to include all the participants in the meeting in the discussion. The most frequent problem regarding the meeting leader responsibilities is that the meeting leaders often "forget" (leave) someone out of the discussion. Most often it is a person (or persons) who are participating in the distributed meeting alone, without collocated colleagues. In order to solve this problem the meeting leader has to ask all the participants about their opinion any time when a discussion is going. He has to create the balance in the discussion. He has to include all the participants in the discussion. He has to ask all the participants to introduce themselves in the beginning of the meeting. He has to follow the meeting agenda. He has to be explicit. He has to take care that the meeting starts and ends in the defined time. He has the task to introduce what the rules (meeting protocol / etiquette) in the meeting are and to control that they are followed by all the participants. *"The minutes that are exchanged by the meeting leader are useful, but it is up to the meeting leader whether he would do it. "*

Another researcher shared his experiences with regard to the role of the meeting leaders when he has been asked the question *"Do you think that the meeting leaders should introduce a type of rules for audio conference meetings?"*

"Yes, I think so, definitely! It is not apparent when only few people are present, I mean when they are only 2 or 3 or perhaps 4. But when you start to have several locations, more than 2 locations it becomes really important, because it can turns into that there are people who want to say something but they are never getting the chance to say something because they are not saying "Hey, I want to say something." And when they are not visible it is important. So, I can remember people having being quite offended by feeling that they have been forgotten. Because in particularly when there are a lot of people in the room and there is another guy on the phone, sometimes it is really disaster, not functioning at all. A little bit depended on the guy on the phone, but very often the entire meeting tends to forget the other on the phone."

Those and similar issues are discussed by Heeren & Lewis (1997) who have found that *"failure to be explicit can lead to feelings of exclusion from decision making and lower levels of personal commitment to the collaborative goals"* (Heeren & Lewis, 1997: 95).

Yankelovich et al. (2004) had also identified that a major behavior problem (in this study identified as a socio cultural problem) is that the *"people gathered in a conference rooms generally ignoring remote participants"* (Yankelovich, 2004: 421). The same study discusses the relationship between the meeting size and the meeting effectiveness: *"The larger the meeting size, the lower the meeting effectiveness. More meeting problems were reported with meetings of 11 or more participants than those with fewer participants"* (Yankelovich, 2004: 421). Then the need of rules (meeting protocol) when the meeting size is larger and when there are more locations included in the meeting is essential. From this it could be concluded that the meeting protocol is applicable mostly to situations with more then 11 participants or a big number of locations. In these situations the meeting leader has to introduce explicitly the meeting etiquette in order to enhance the meeting effectiveness.

6.4.3 Language issues

The idea that audio conferencing is not the perfect communication tool was the generally shared opinion by the respondents from the interviews. The first argument that easily came to their mind was that audio conferencing does not provide rich enough channels for exchanging information and that they would like to have visual contact with their interlocutor(s) in order to easily solve the language problems. Albert Mehrabian (1972) has found that only 45% of the emotional meaning of a message is communicated by a person's spoken voice. The remaining 55% of a message is communicated through non-verbal cues including gestures, posture, and facial expressions.

The communication in a foreign language is a challenge, shared one of the researchers and added that when it is mediated by a telephone it could turn into a disaster.

"First when the people are not so good in English there are difficulties in the communication. Second when the people have an accent but you communicate with them face-to-face, then you can see their facial expression, body language and they can also repeat if you do not understand and try to readapt. But when the conversation is through teleconference, then you do not see the person face-to-face, they can not realize that you have a lot of problems to understand what they are saying."

One of the limitations while communicating with a stranger through the mediation of audio conferencing is that the communication is mainly verbal. The extent to which we are good at communicating verbally is purely personal characteristics comprising diverse abilities, which could be divided into two main groups: first, to express ourselves orally and second, to hear, comprehend and interpret the sound information expressed by others. These two groups of abilities are developed in a different degree in any single individual and the combination of those two forms four character groups which could be generalized as follows: a person is good at expressing verbal information and he is good at interpreting verbal information, second – he is good at expressing verbal information, but he is not good at interpreting verbal information, third character – he is not good at expressing verbal information, but he is good at interpreting verbal information, fourth character – he is not good at expressing verbal information and he is not good at interpreting verbal information. Any of these four characteristics are not applicable for a person taken out of a context, but they are the gradations where one person is compared to another or others. We can not say that one is always good at expressing verbal information, because in different conditions he could be good or not good. For example, if a native English without accent is very good at communicating and interpreting English speech in England then the extend to which he is good in communicating with native English Americans is different. An interviewee states that: *"the Indian are very good in English, but when they speak*

English it is very difficult for me to understand them, due to their accent." In the case described from one of the participants in the interviews the Indians have excellent verbal abilities for communication, but only in their own environment. In other conditions (of different states, countries or continents), where the abilities of the people for verbal communication in English could be also very good the abilities of the Indians are evaluated as poor. Thereby the following conclusion is drawn: in the cases of group work the work tasks is preferable to be distributed according to these 4 categories. Why?

The ideal situation is when there is a balance between the people who are good at expressing verbal information and those who are good at interpreting this information, because those who like to speak more, they do not have time to consider in depth the discussed issues. The deep analysis could be made by the people who are good in interpreting information. Through their comments and remarks they lead the discussion in the right directions and ensure an efficient outcome of the meeting.

If the number of people who have good verbal skills for expressing their meaning (but not good at listening) prevail in the meeting, then there is a lack of people that could interpret the information and take a decision. Then the outcome of the meeting is poor - the task is rarely well solved and the time is used inefficiently in discussions.

The other situation is when there is a prevalence of people with good interpretive skills; then the outcome of the meeting is as well inadequate. An interviewee shared:

"Communicating in foreign language is a challenge. The people in Ukraine are not very good in speaking English, at least they do not think so, they feel very uncertain when they speak English. Usually the language makes it harder in audio conferencing. But I believe that the people speak les if they have to use English."

In this case the efficiency of the meeting could be increased if the participants are prepared with presentations (or similar) in advance. But as a whole it is missing the discussion part of the meeting and thereby important decisions could not be taken. Another interviewee said that in the communication with Ukraine there is a mutual

agreement for not using audio conferencing because of the language problem. Both sides have decided to meet face-to-face, on long workshops, for the important discussions, because only in the face-to-face meetings the language problem is manageable. They are using audio conferencing only for short and simple conversations, and emails for exchange of important information.

There is no fourth case constituting only of "ignorant" participants, because as it was explained before these categories are comparative, they are not universal and complete, because there is always differences and gradation of the people's skills in a group, and the skills are examined here as comparative and complementary to the skills of the rest of the people in the group.

Overcoming the language problem

The five main approaches used by the employees in R&I for solving the language problems are: ask to repeat, get used to the accent or dialect, use after-call email for clarifying the points discussed during the audio conference, replace the audio conference communication by email communication, don't use audio conferencing. There was another meaning that it is difficult to do something about solving the language and the accent problem and it could be done something about improving the technical side in order to improve the meeting effectiveness.

6.5 Conclusion

The main themes from the interviews were extracted and ordered in three different categories – technological, socio-cultural and linguistic. In addition the results from the interview study were analyzed and compared with similar findings discussed in the research literature in order to shape a comprehensive view of the questions of present interest.

The thoughts of one of the researchers about the most frequent problems in audio conferences summarize the interview study findings. *"The most often problems in*

audio conferences are due to: first the quality of the sound, the second major problem is the background noise and the third issue is the common language, and also quality of the devices."

7. Observation study

This chapter consists of the following parts: aim, participants, method, results and discussion. In the first part the aim of the observation study is explained. The second part describes who the participants in the current study are. The third part explains the characteristics of an observation as a research method. The next part presents the results in the observation study. In the last part of this chapter a discussion of the results is presented.

7.1 Aim of the study

The primary aim of this study was participant observations to be conducted in Telenor R&I, because the participant observation has the potential to uncover contextualized, truthful data. But regarding the limitations described in 5.5 this aim could not be realized. The observation study aimed to get a comprehensive picture of the situations occurring during audio conferencing sessions. Considering the fact that the conducted two observations were the first experienced audio conferencing session for the author, the current study is regarded as a valuable contribution to the whole investigating process.

7.2 Participants

Participants in the observation study were people from different countries working together distantly on a specific project/s financed by ETSI. All of them have rich background experience as users of audio conferencing and some of them not only as users, but also as experts knowing the fine points of the audio conferencing became clear from the comments in the short chat with some of them. It is interesting to be noted that they were having different mother tongues and all of them were men; the conversation was in English. There were 6 distributed participants on the first

meeting and 10 participants working collaboratively on EG 202670 project on the second meeting.

7.3 Method

The basic approach in the ethnographic methodology is observation. The observation endeavour is conducted principally through note taking during the observation process. The aim is to collect detailed, specific, descriptions such as examples of speech, ways of speaking, initiating conversation, beginning and ending an event, timelines and any other concrete, physical detail of behavior, speech, or the environment. Considering that audio conferencing sessions were observed distantly (not participating in a group of attendees) the following data typically collected in the ethnographic observations were not possible to be observed: body language, gestures, the physical environment (the room, facilities), social environment (the community this group is a part of). The notes taken during the second observed meeting could be seen in Appendix B; the reflections of the meeting are not included.

The second step of the observation method is writing a report based on the notes taken from the field – data have to be ordered and classified. The third step is interpretation of the data - patterns should be sought and analyzed.

7.4 Results

It has been observed two audio conference meetings of an ETSI project. The audio conferencing system that was used for the meetings from all of the participants was a VoIP solution called GoToMeeting[10]. The first meeting called officially HF G2M VoIP Test was held on Wednesday 15th July 2009 from 14:00 to 14:30 Romance Standard Time. The second meeting was HF (human factors) discussion of EG

[10] http://www.gotomeeting.com/fec/

202670 / HF#49_17r1; it was held on Thursday 16th July 2009 from 14:00 to 15:30 Romance Standard Time.

The participants were connected to the audio conferences through their computers. The use of headset was recommended. It has been also reserved an audio bridge as a backup solution. On the both meetings there was a chairman who chaired the meeting. On the second meeting there was an agenda provided in advance in the reminding for the meeting email. The audio quality during both observed sessions was satisfactory good.

7.4.1 Visibility who is in the meeting

Before one enters in the virtual meeting room, the system shows a window where one can enter her name and additional data, for example the company or the country she is representing. This data is shown in a list of all the participants during her presence in the virtual meeting room.

In the beginning of the conference, when a person "enters" in the virtual meeting room a recorded welcome message is informing every new participant that s/he is present in the meeting and invites her to introduce herself to the other participants.

7.4.2 Mute button

In the beginning of the second meeting the chairman of the meeting asked all the participants to mute their microphones while not talking. During the conversation it was noticed that none of the eighth (8) attendants muted their microphones with little exceptions for a short time. During the first meeting the situation was absolutely the same.

7.4.3 Background noise

During the second meeting there was someone who was typing on the computer. The reproduced noise entered in the communication channel as a background noise which

was not so loud and thereby not so much annoying, but nevertheless, it was perceived as an obstacle.

7.4.4 Volume control

What made an impression was that there were very sharp differences between the levels of loudness of the voices of the different participants in the distributed meeting. Some of them could be heard very loud while one of them could be barely heard.

7.4.5 Accented speech

There was one of the participants, who had a strong French accent, but he spoke very slowly and it made the speech understandable.

7.4.6 Text messages supporting audio conference

During the first meeting there was one participant who could not talk in the conference, but he was able to listen to the discussion and participate in it by typing messages.

7.4.7 Identify and show who is speaking

There was an interesting situation in which several 2, 3 or 4 participants are discussing something in the beginning of the main conversation while waiting all the participants to enter in the virtual meeting room. Then another participant entered and began to adjust his microphone. The system indicated him and showed his name in the list of the people who are speaking or "creating noise". At the same time everybody was able to hear that someone was breathing heavily in the microphone and all were wondering who the person is. Some of the participants mentioned it. Soon all of them decided that it was the new participant who creates this annoying sound, because the system indicated a noise entering through his microphone. He was not sure, because he was not able to know what his microphone is picking up. He was

asked to take away the microphone. After a while it turned out that there was another person creator of the heavy breathing.

7.4.8 Overlapping

During both observed meetings speech overlapping occurred several times, which was interpreted as a common situation.

7.5 Discussion of the results

The system GoToMeeting that was used for the observations is very similar to Rendezvous system described in (Kellogg, 2006). Rendezvous is also a VoIP based audio conferencing system which provides extra information about the conference participants. It is not surprising that the results from the current observations are also very similar to the findings published in Kellogg et al. Both systems are using the idea of "social proxy" introduced in (Erickson, 1999) of creating social translucence as *"providing perceptual cues that lead to awareness and accountability"* (Erickson, 2000). GoToMeeting and the IBM's Rendezvous *"provide a shared, minimalist visual representation of people and their activities in an online environment"* (Kellogg, 2006). The results from the survey made with the users of Rendezvous have shown that the system *"enhance the user experience of conference calling and improve meeting effectiveness"* (Kellogg, 2006).

GoToMeeting system was found to be user friendly, for the following several reasons: It acts as a call attendant - it welcomes the participants and invites them to introduce themselves, can identify and show who is speaking.

One of the disadvantages of the system is that it captures all the sounds, not only the speech and indicates and reports them equivalently. It does not have an integrated mechanism through which to differentiate what is speech and what is not speech. The system constantly indicates whose microphones are receiving sound or noise and shows the participants' names on the screen. It is possible that more than one person

are creating noise or speaking at the same time, and then the system shows all their names. For a listener it is not understandable who is the person/s speaking and who is creating noise. Thereby the system should have a mechanism through which to identify and distinguish speech from noise.

The differences of the levels of loudness of the voices of the participants in the distributed meeting created discomfort to the listener and is a characteristic which drifts the user experience far from the reality. It is assumed that the reasons for the poor volume control have diverse nature as for example the different distances to which the participants are situated to their microphones, the different quality of the devices and the settings of the software that any of the participants is using. Thereby the system should have means for equalization of the loudness level.

8. Questionnaire study

The next research method used in the entire study in order to perform *methodological triangulation* is survey or questionnaire study. The questionnaire study is used in conjunction with the interview and observation studies to verify and clarify the results and to deepen the understanding gained from the field. The various data collection methods have different advantages and disadvantages and the area where one method is weak there another method could be applied. For example one can not measure user satisfaction through observation or interview, but only through a quantitative survey.

This chapter is structured as follows: the chapter begins with an introduction and describes the aim of the questionnaire study; then the chapter continues with an explanation of the methods used for conducting the study and the making of the questionnaire; in the next paragraph the participants in the study are described; the chapter ends with a presentation of the results form the conducted survey together with a discussion of the results.

8.1 Aim of the questionnaire study

The goals of the questionnaire study are directed by the findings from the conducted before qualitative interviews. The results from the interviews helped to become acquainted with the kind of quantitative data it is possible to get from the investigated domain. The questions asked in the questionnaire, and those used in the interviews are similar, but those in the questionnaire are more specific, because the interviewer is not available to explain or clarify any ambiguities to the respondents. Therefore an objective during the preparation of the survey was: questions to be well formulated and brief, with only one clear meaning where ambiguity should be avoided. This helps to the respondent by making their efforts efficient and time consistent.

In addition, the results from the survey provide the opportunity for quantifying some factors that could not be measured through qualitative interviews and observations, like for example effectiveness, efficiency, satisfaction and other.

Goals of the questionnaire study:

- Exploring user attitude towards the audio conferencing systems;
- Quantifying user opinion about the factors influencing the quality of experience in AC;
- Measuring the user satisfaction from the audio conferencing systems.

8.2 Method – describtion of the questionnaire study

The questionnaire study was conducted in practice by using an online web based survey. Before the web-based questionnaire was made a paper-based version was prepared. Making a questionnaire is a creative work requiring knowledge and imagination. It is of great significance to know how to ask the right questions. The imagination supports creativity by making the entire mind to work for realising the goal in its best form.

The survey method itself has advantages in comparison to the other ethnographic methods employed in the present study – it provides response from a large representative sample in a short amount of time. But the true strength of the survey method consists in the possibility to combine it with the other two methods that are used here because it supports the clarifying of the findings from them. Also the online web based surveys have many advantages compared to the paper based surveys. First and foremost they offer faster response than the paper based surveys and can transfer the data automatically into the database. Another advantage is that they are interactive and provide immediate data validation and can impose the selection of only one or several responses.

The data collected through survey is subjective. It represents the subjective user opinion about the audio conferencing systems in Telenor. Also the method through which the survey is conducted is not as reliable as if the survey was conducted exactly after several consecutive audio conferencing sessions with the aim to measure the quality of experience in particular sessions. In the current case the user *is "open to primacy and recency memory effects"* (Watson & Sasse, 1998).

8.2.1 Design of the content of the survey

The way the information is structured on the screen can influence how the people answer the questions. The sequence of the questions is also of a great importance and can affect the answers. Thereby it was appropriate to order the questions with respect to respondents' logical thinking. The questions with similar themes were grouped together into related topics to make it easier, encouraging and more logical to complete the questionnaire. All the questions built the body of the questionnaire and thereby the main idea of the questionnaire had to unify all the questions. In addition the questionnaire consists of all necessary questions, in order for it to be functional.

The questionnaire (see Appendix D) consists of an introduction and the following groups of questions: General Language questions; Communication: Purpose and language, Language, Audio conferencing system, Challenges in Audio conferencing, Audio quality issues, Socio-cultural issues; Demographics and Final questions.

Introduction

The survey starts with an introduction which encourages and motivates the respondents to complete the questionnaire. The fact that the results of the survey will be used for improving the quality of user experience of audio conferencing should be motivating for the workers in Telenor to fill in the survey. This idea was confirmed by the following sentence:

"If you are surveying members of an organization, the members may be more likely to respond if they think the organization is asking their opinions on how it can best

meet their needs. The same could be true when you are surveying users of a particular service." (Creative Research Systems, 2009)

In the introduction part of the survey there is information about the purpose of the questionnaire and about the pre-estimate time that it takes to complete the survey. It is explained how the collected information will be used. The confidence is a very important aspect when conducting a survey. The respondents' curiosity could be intrigued, but if they do not have trust, they will be likely to skip sensitive questions. Sensitive questions are all those questions concerning respondents' skills and competences with regard to their work. It is expected that nobody wants to admit that one lacks some competences to do their job. Thereby, the respondents are assured in the introduction part that the survey is anonymous. Also by creating confidence it is expected that the response rate to the more sensitive questions, investigating users' competences will be higher.

Personal questions / Demographics

At the beginning of the survey design, it was found appropriate to set the personal questions at the very beginning of the survey because *"many questionnaires start by asking for basic demographics information…"* (Rogers & Sharp, 2007: 310). It was considered that it is natural to start with the personal questions, like age, gender, experience and occupation, because it can make the filling of the questionnaire encouraging. It was planed to place the difficult questions, those that require specific answers and demand mental efforts or recalling, afterwards. But after that it was realized that by asking about personal information at the very beginning does not create trust, and thereby it is not suitable. Thereby the personal questions were placed at the end of the survey. Iraossi (2006) confirms these thoughts:

„Questions like demographics or personal information are usually best to introduce towards the end of the survey. This way, respondents are likely to have already developed confidence in the survey's objective" (Iraossi, 2006: 74-78).

Because the questionnaire is anonymous the demographics will bring the personal characteristics of the respondent. The *"background information is useful for putting the questionnaire responses into context. For example, if two responses conflict, these different perspectives may be due to their level of experience…"* (Rogers & Sharp, 2007: 310). In the analysis this type of data was used for giving specifications when generalizing the results. There could be found similarities in the responses when they are sorted by age, gender, occupation and experience. For example it could be taken into consideration the relationship between the age and the perceived audio quality where the respondents age 51-65 are possible to have (more often than the other age groups) problems with audio quality that could be due to ear. Another group could be formed by the novice users, those who have worked less than a year at Telenor. It is presumed that they are not so good at using the systems and setting up a conference call.

8.2.2 Design of the questions and the answers in the survey

The following several paragraphs describes the design choices made for the questions type and the answers type and in addition some examples of these design choices from the questionnaire are provided.

Allowing answers "I do not know" and "Other"

During the testing process of the survey the need of using the answers "I do not know" and "Other" become evident. The answers "I do not know" are wasted, because they do not contribute to the future analysis, but though they were included as an option in the range of answers because some respondents could feel frustrated without such a choice. And also because the *"respondents who feel they are being coerced into giving an answer they do not want to give often do not complete the questionnaire"* (Creative Research Systems, 2009).

The optional answer “Other” is used together with the invitation “Please specify!” in order to collect detailed picture of the respondents’ opinion, in case their choice is not among the choices that appear.

Answer choice types

There are several different types of answer choices that have been used in the survey: multiple choice, numeric open end and text open end (sometimes called "verbatims"). *“Rating Scales and Agreement Scales are two common types of questions that some researchers treat as multiple choice questions and others treat as numeric open end questions”* (Creative Research Systems, 2009).

An example from the survey about *Text open answer choice/ verbatims* is the following question, demanding typing of text:

1. *Do you have any other recommendations for improving the quality of audio conferencing?*

    1.1. No

    1.2. Yes

        1.2.1. Please describe!

Rating Scales

Scales represent frameworks through which one can measure and quantify ones attitude or opinion towards something. *„The purpose of these is to elicit a range of responses to a question that can be compared across respondents”* (Rogers & Sharp, 2007: 313).

This is a *rating scale* (or more concrete *likert scale*) example from the survey:

☐ Daily

☐ Weekly

☐ Monthly

☐ More rare

☐ Other. Please specify!

Figure 9: Likert rating scale

"Likert scales rely on identifying a set of statements representing a range of possible opinions, while semantic differential scales rely on choosing pairs of words that represent the range of possible opinions" (Rogers & Sharp, 2007: 313).

This is an example of a semantic scale. Semantic scales *"explore a range of bipolar attitude about a particular item. Each pair of attitudes is represented as a pair of adjectives"* (Rogers & Sharp, 2007:314).

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Figure 10: Semantic scale; 1 is extremely not important, 7 is extremely important

Researchers are not unanimous on how many points are needed in the scale and there are different arguments for and against the different meanings. *"Many surveys use a ten-point scale, but there is considerable evidence to suggest that anything over a five point scale is irrelevant. This depends partially on education. Among university graduates a ten point scale will work well. Among people with less than a high school education five points is sufficient. In third world countries, a three-point scale (good/acceptable/bad) may be all some respondents can understand"* (Creative Research Systems, 2009).

Another aspect in regard to the scale points that was considered is whether the scale has to have an odd or even number of the points. The even number scale forces the respondents to make either the one or the other decision, where the odd number scale gives opportunity to choose the medium. And since *"respondents tend to select values nearer the center"* (Roger & Sharp, 2007: 315) and it is understandable that it could be their most honest decision, it was decided not to make the interviewees feel frustrated when forcing them to make a decision.

Based on the aforementioned arguments it was decided to use the scales from 1 to 5 or 1 to 7 in the survey because of the level of easiness for estimation and evaluation of the personal opinion in accordance to the scales.

8.2.3 Testing the questionnaire and piloting the survey

The questionnaire and the survey were thoroughly piloted in four stages: first the questionnaire was reviewed by knowledgeable analysts; second, typical participants completed the questionnaire by using a think-aloud protocol; third, several survey versions were reviewed and changed until the final version was found appropriate; a final check to catch small errors and to estimate the time for completing it were conducted.

The questionnaire was tested with the traditional paper and pencil method by 4 persons where 2 of them made it twice after several changes were made. The questionnaire was also edited several times following the advices of 5 knowledgeable analysts. The process of editing transformed the survey from a sequence of messy and not so clear questions to a structured and goal oriented questionnaire. The test revealed unanticipated problems with question wording and instructions and helped to see if the interviewees understand the questions and give useful answers.

8.2.4 QuestBack

QuestBack is the system that was used for generating interactive online survey. The same system was also used for sending email invitation and remainders to respondents. QuestBack automatically creates analyses from the received responses in the forms of graphs and tables, where there are integrated options for choosing which data to be compared. The system could also export the results in the common formats like .pdf, .xls, etc and could send follow up anonymous messages to the respondents.

8.3 Participants

An email-invitation with a request to fill in a survey together with a link to the survey was sent to 175 persons in Telenor at Fornebu. The email was sent on 22 November and a second reminder was sent on 1st December. The survey was active 18 days, until 08 December and was filled in by 67 respondents. The response rate was 38%.

Everyone chosen to be invited to fill in the survey was from Telenor Research & Innovation because the employees there are using AC more often then other departments. This department is supposed to have a lot of communication with other countries and thereby they are expected to be one of the most experienced people with regard to the audio conferencing in EFL. The invited respondents have diverse positions and roles in the company - researchers, advisors, secretaries and managers.

8.3.1 Target population and Samples

There are two main components in determining who will be interviewed. The first is deciding what kind of people to be interviewed. This group is often called the *target population*. The next thing to decide was how many people have to be interviewed, which is called the *sample*.

The main characteristic of the target population is that all the people are employees of Telenor and are working in Fornebu. These are the people of main interest for the current study. The survey is designed only for users of the audio conferencing systems in Telenor (in Fornebu), because the questions in the survey are consistent with the pre-conducted face-to-face interviews with the employees working in the research department in Telenor, Fornebu.

The second step was to decide about the sample - how many people have to be interviewed in order to get an objective image of the targeted population's attitude towards the audio conferencing and to be able to quantify their opinion about the factors influencing the quality of experience. If the sample is small then it could not reflect precisely the targeted group. But if it is too large then the precision is not

symmetrically big. *"For example, to increase a sample from 250 to 1,000 only doubles the precision"* (Creative Research Systems, 2009). This relationship is not linear.

The number of employees in Telenor, Fornebu is about 5500. The confidence level used for the calculations is chosen to be 95 %, because it is the most often used by the researchers. It shows that one can be 95% sure about the authenticity of the received results. Thereby it was decided that a sample from 175 persons is enough to represent the targeted population's opinion. By using a formula and the aforementioned numbers it was calculated that the confidence interval is seven (7). It shows that if 60% of the respondents have picked an answer, then it is expected that 53% to 67% of the whole population would pick the same answer.

8.4 Results and Discussion

AT is usually used for qualitative analysis where the data that is object of the analysis is collected through ethnographic methods. But as it was mentioned in chapter 3.3, AT provides a very good structured diagram representing the relationship between the different components and with very useful and practical naming of the components. This advantage of the theory makes it useful tool for analyzing both qualitative and quantitative data either in combination or individually. The approach that is undertaken here is firstly analyzing systematically the quantitative data derived from the survey according to the categories provided by the AT's theoretical framework and the developed approach in the current study represented by the AT-matrix (Table 1); and secondly supporting and enriching the analysis by additional qualitative description derived from the interviews and observations.

8.4.1 Activity theory and activity system

Definition of Activity: *"activity refers to a specific level of subject-object interaction, the level at which the object has the status of a motive. The motive is an*

object that meets a certain need of the subject" (Kaptelinin & Nardi, 2006). The motive in the current case (in a broad sense) is the communication or exchange of information, whereas the motive in a narrow sense could be specific for the company's own policy and strategy. The individual subjects could also have their own motive. The subject is a group of people distributed over different locations. *"Activity ... is a unit of subject-object interaction* [distributed group-communication] *defined by the subject's* [the distributed group's] *motive. It is a system of processes oriented toward the motive* [communication]*, where the meaning of any individual component of the system is determined by its role in attaining the motive"* (Kaptelinin & Nardi, 2006). On the basis of this definition of activity the investigation could be directed toward the question "What are the processes accompanying the activity". In the following subchapters the different factors and processes accompanying the activity are ordered in subgroups. Those subgroups are the main components of AS.

An activity system is created any time when a subject (an individual or a collective) has the objective to conduct an audio conferencing session (in Telenor). The subject has concrete motive which directs him toward the object. The motive is usually related to the organization's strategy and plans. This main activity is divided (abstractly and for analytical purposes) into actions. For instance, if the main activity is the audio conferencing, then an action motivated from the main activity could be the selection of technology which to mediate the communication process. The individual actions could not be understood separately out of the context of the collective activity, because each separate action is motivated by a main activity. Each of the actions is goal oriented and the goals are achieved by fulfilled actions. The actions are also divided for analytical purposes into operations. The operations are routine, non-conscious actions. The activity is divided into actions and operations according to the following scheme:

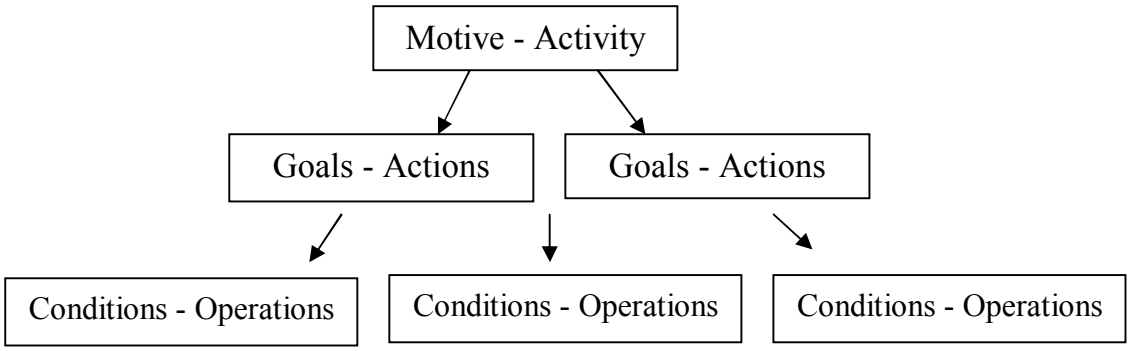


Figure 11: Hierarchical structure of activity

8.4.2 Activity = Subject – Object interaction

Frequency of activity

The main characteristics of the activity are represented by the relationship Subject-Object. The activity itself is a result of the Subject-Object interaction. This relationship gives an answer to the question how often the activity occurs (Figure 12): 4.5% of the respondents are using audio conferencing at Telenor daily, 40.3% - weekly, 28.4% - monthly, 25.4% - more rarely. It makes an impression that the multitude of people is using AC at least ones per week. But there are a big number of respondents that are using it more rarely, which would mean that they are not very much experience with AC and not so good acquainted with the related issues. Despite this their sincere answers are valuable for the study. The survey gives options for choosing "I don't know" when answering questions, which eliminates the possibility for giving wrong answer and ensures reliability.

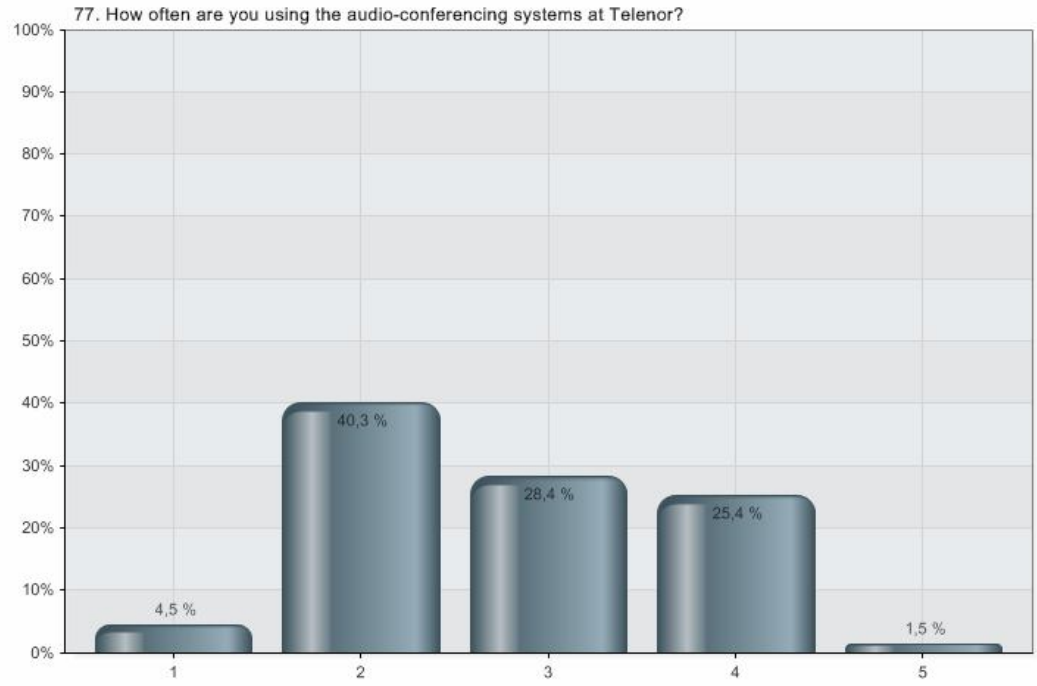


Figure 12:"How often are you using the audio conferencing systems at Telenor?"

EFL

Another issue that the relationship Subject-Object represents is the answer of the question "How often are you using English for audio conferencing". To this question 6% have answered that they are using it daily, 31.3% - a few times per week, 49.3% - occasionally (a few times per month) and 13.4% - rarely.

56.5% of the respondents answered with 'yes' when they were confronted with the question "During an audio conference have you communicated with anyone whose English you can not understand?". 58.3% of those who answered with 'yes' have experienced it rarely, 38.9% - occasionally and 2.8% - often. There are many factors affecting the speech intelligibility (discussed in section 2.2.3) as for example the accent, echo / reverberation, softness of the speech, limited bandwidth (discussed in section 2.2.1) and not least the level of command of the language (discussed in 6.4.3).

Nature of the problems in AC

The respondents were asked to point out the most frequent problems experienced in audio conferencing, which is the general research question in the current study. This issue was in addition widely explored through variety of questions, in order to deepen the understanding of the user experience with audio conferencing.

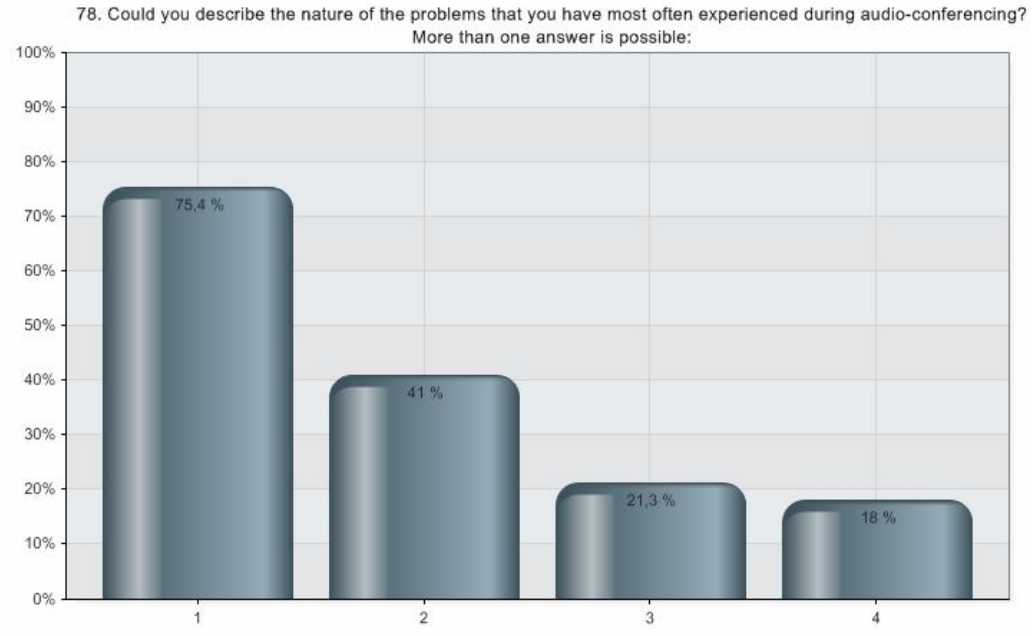


Figure 13: The nature of the problems experienced during audio conferencing

The results have shown that the most frequently encountered problem during audio conferencing has a technical nature, where 75.4% of the respondents have agreed upon it (see Figure 13: The nature of the problems experienced during audio conferencing). Language problems are experiencing 41% of the users and according to 21.3% of them the socio-cultural issue is also occurring as problem during audio conferencing. There is an additional group of people represented by 18% which have experienced other problems during audio conferencing. It was identified that the most of the described problems under the category "Other problems" are part of the three aforementioned categories – technical, language and socio-cultural. The problems that the respondents have pointed are:

- Three respondents have answered with *"No problems"*.

- "*Voice control* (discussed in 7.4.4), *feedback*". The voice control is a technical issue, which was reported in the observation study. There are audio conference applications on the market with newer and improved features for automated voice control which could be a good solution for solving the voice control problem.

The feedback is a major disadvantage of the audio conferencing activity itself. Endeavors by various research groups to solve this issue have resulted in creation of VoIP applications with integrated features for sending feedback. For example trough Rendezvous application, described by Kellog (2006) could be sent non-interruptive feedbacks to the person speaking or to other participants in the distributed meeting.

- *"If [there are] too many participants in one room, the conversations/discussions are difficult to comprehend for users on the phone."* This problem has a dual nature – socio-cultural and technical. Yankelovich et al. (2004) is arguing that there exist a correlation between the meeting size and the meeting effectiveness: *"The larger the meeting size, the lower the meeting effectiveness"* (Yankelovich, 2004). The same study reports that the problems exist most often in meetings with more than 11 participants.

- *"Network congestion in Asian countries"* and *"Poor quality of connection with some Asian countries"* are similar technical problems. Measures about the communication with Asian countries are presented in Table 3: The mean values for Audio quality, EFL, Meeting protocol for 12 countries; 1 is extremely low quality, 7 is extremely high.

- *"Too poor audio quality"*. The audio quality is the major technical issue discussed in 8.4.4.

- *"Waiting for connecting participants"* and *"People entering conference calls too late"* could be referred as an important socio-cultural problem. This and other behavior issues are discussed in the next paragraph.

Behavioral patterns reducing QoE in AC

Behavioural patterns are usually looked as occurring in a subject – subject interaction, but they are placed in the current Subject – Object interaction section, because the element Subject in AT unifies all the subjects involved in the activity. It is assumed that their common Object is communication. Thereby it is argued that the behavioural patterns occur in the process of Subject – Object interaction.

The respondents were asked to rate 17 different descriptions of distracting behaviour using a scale from 1 to 7, where 1 is "not distractive" and 7 is "extremely distractive". They were asked the following: "From the following list please rate the ones that you feel distract or reduce the quality of your audio conferencing meetings". Speaking on another phone during AC is rated as the behaviour reducing in a highest degree QoE in AC. Other behavioural patterns rated as very much distractive are overlapping, attending the conference from public places, not understanding the way of speaking English, coming late to the meeting, speaking too quietly/ softly. Table 2 shows the behavioural patterns reducing QoE in descendent order. Figure 14 shows a graphical representation of the same results.

Behavioral patterns	Mean value
Speaking on another phone during a meeting	5,54
Overlapping /"talking over" the conversation	4,9
Attending the conference from public places	4,89
You do not understand her/his way of speaking English	4,84
Coming late to the meeting	4,63
Speaking too quietly/ softly	4,59
She/ He can not understand your English	4,13
Speaking too fast	4,1
Breathing in the microphone	4,03
Interrupting	4
Whispering with colleagues	3,95
Hearing the electrical sound from the near-by devices	3,89
Typing on the computer	3,71
Drinking or eating	3,29
Speaking too loudly	3,24
Moving doors, windows, chairs	3,16
Speaking too slowly	3,03

Table 2: Behavioural patterns reducing QoE in AC

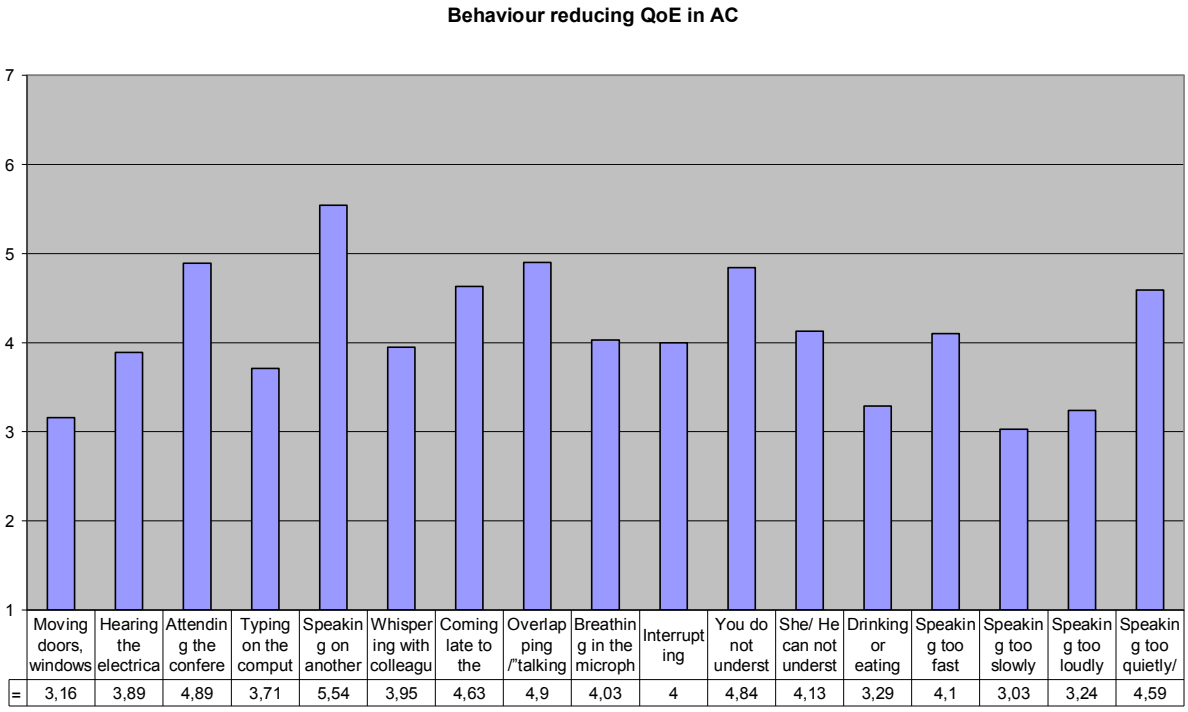


Figure 14: Behavioural patterns reducing QoE in AC

The degree of distraction is an important measurement when looking at the behaviour patterns reducing QoE. In addition to this the frequency with which behaviour patterns occur is another key measurement. The frequency of the pattern 'late start of AC' was investigated. The results are presented in the next section.

Late start of AC

82% of the respondents have experienced late start of a scheduled audio conference meeting. According to 75% of them the reason for late start is that the "people are coming late for the meeting", and 74% think that it is due to "technical difficulties in setting up the conference". The relationship between the lack of knowledge for setting up AC (discussed in 6.4.1) and late start of AC is obvious and it affects QoE. The words of a researcher from Telenor R&I illustrate this issue:

"We were a group of people and had to have a teleconference meeting. But no one knew how to set up the teleconference. We spent about 20 minutes until we connected to the meeting. But then the purpose of our participation in the meeting was not

fulfilled. And the quality of the meeting dropped, because we were not able to discus the issues that the other participants were discussing."

8.4.3 Subject

The typical user

The Subject of the activity system is any individual or groups of individuals involved directly in an audio conferencing session, e.g. taking part in the activity. But who is the "typical Subject" in the current activity system? The answers of the questions in the first part of the survey can describe the typical respondent:

The first language was Norwegian for 85% of the respondents which most probably means that those people are Norwegians. Another clearly distinguished group is formed by the rest 13% of the respondents whose first language is English. For 2% of the respondents the first language is neither Norwegian nor English. By using the answers in the log it is possible to investigate the differences in the opinion between the first two groups regarding the language group questions, e.g. how the EFL is perceived by the native English speakers and by the rest of the respondents.

Culture and language

This data from the survey represents a very important aspect of the investigation called 'culture' because a main starting point of AT is the philosophical idea about the social nature of the human mind. *"Human beings are shaped by culture, their minds are deeply influenced by language…On the other hand, the world itself is fundamentally social. The entities people are dealing with are mainly other people and artifacts developed in culture"* (Kaptelinin & Nardi ; 2006, p.37). A similar idea is inferred in the works of Edward Sapir and Benjamin Lee Whorf, also known as the *Sapir – Wholf hypothesis. "In its most extreme version 'the Sapir-Whorf hypothesis' can be described as consisting of two associated principles. According to the first, linguistic determinism, our thinking is determined by language. According to the*

second, linguistic relativity, people who speak different languages perceive and think about the world quite differently" (Chandler, 1994). Sapir (1929) argue that:

"Human beings do not live in the objective world alone, nor alone in the world of social activity as ordinarily understood, but are very much at the mercy of the particular language which has become the medium of expression for their society. It is quite an illusion to imagine that one adjusts to reality essentially without the use of language and that language is merely an incidental means of solving specific problems of communication or reflection. The fact of the matter is that the 'real world' is to a large extent unconsciously built upon the language habits of the group. No two languages are ever sufficiently similar to be considered as representing the same social reality. The worlds in which different societies live are distinct worlds, not merely the same world with different labels attached... We see and hear and otherwise experience very largely as we do because the language habits of our community predispose certain choices of interpretation." (Sapir 1958 [1929], p. 69)

It can be argued that the similar languages have similar cultural constructs, because of the principle of linguistic determinism which states that our thinking is determined by language. In the current case the investigation was made about the communication in English as a foreign language. But although the people are using a second language for communication, the construct of the communication in a foreign language is based on the first language, because *"human beings are shaped by culture, their minds are deeply influenced by language"* (Kaptelinin & Nardi; 2006, p.37). The results from the survey are corroborating this statement: The respondents were asked to evaluate the communication in EFL with different countries. The communication in English with Denmark is evaluated as one of the best among a list with 12 different countries (see Table 3). It gets a mean score of 5.33 in a scale from 1 to 7, where 1 is extremely low and 7 is extremely high. The communication with Sweden in EFL is also evaluated as good – 4.94.

The audio conferencing is a multilateral activity involving diverse participants which characteristics are difficult to be predicted. The meeting effectiveness is strongly depended on the participants' characteristics. By using the data gathered about the different locations with which the participants in the survey are communicating, predictions could be made about the weaknesses of the communication with those countries. It is argued that by having this data and if it is used practically, the attendees in the audio conferencing in Telenor could use the knowledge to predict weaknesses in their communication with remote colleagues, drawing proper strategies about the communication with a particular location and managing the communication. In the following pages it is presented the graphics with the assessments showing the audio conferencing communication characteristics with 12 different countries. The data is presented according to the three main categories – audio quality, communication in EFL, and meeting protocol.

The numbers in show the mean value (Grouped Statistics Data Mean Value in the Statistics) calculated from the ratings given by the respondents using a scale from 1 to 7, where 1 is extremely low and 7 is extremely high. The respondents rated the audio quality, EFL and the meeting protocol in the communication with 12 destinations.

Country	**Audio quality**	**EFL**	**Meeting protocol**
Norway	4,14	4,94	4,87
Denmark	4,54	5,33	5
Sweden	4,85	4,94	4,4
Finland	3	3	3
Hungary	5	4,69	5,4
Montenegro	4,67	4,7	4,7
Serbia	5	5,33	5,5
Ukraine	4,28	4,86	3,67
Pakistan	3,17	4,58	5,27
Bangladesh	3	4,45	5,8
Thailand	4,5	3,93	5,11
Malaysia	4,25	5,04	5,38
India	3,5	5,2	5,33

Table 3: The mean values for Audio quality, EFL, Meeting protocol for 12 countries; 1 is extremely low quality, 7 is extremely high

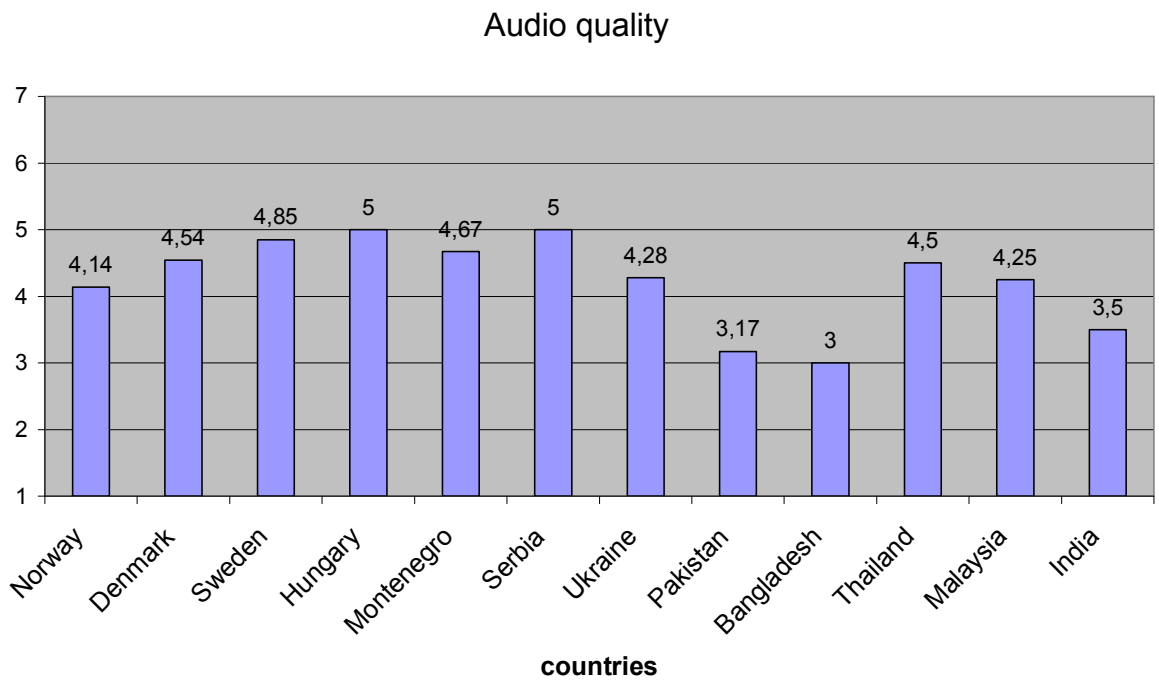


Figure 15: Audio quality for 12 countries; 1 is extremely low audio quality, 7 is extremely high audio quality

The respondents from Telenor are assessing the audio quality of the connection higher than the average (higher than 4) when they are communicating with Denmark, Sweden, Hungary, Montenegro and Serbia; they are assessing the audio quality of the communication with Pakistan, Bangladesh and India lower than the average; and the audio quality with Norway, Ukraine, Thailand and Malaysia is assessed as average.

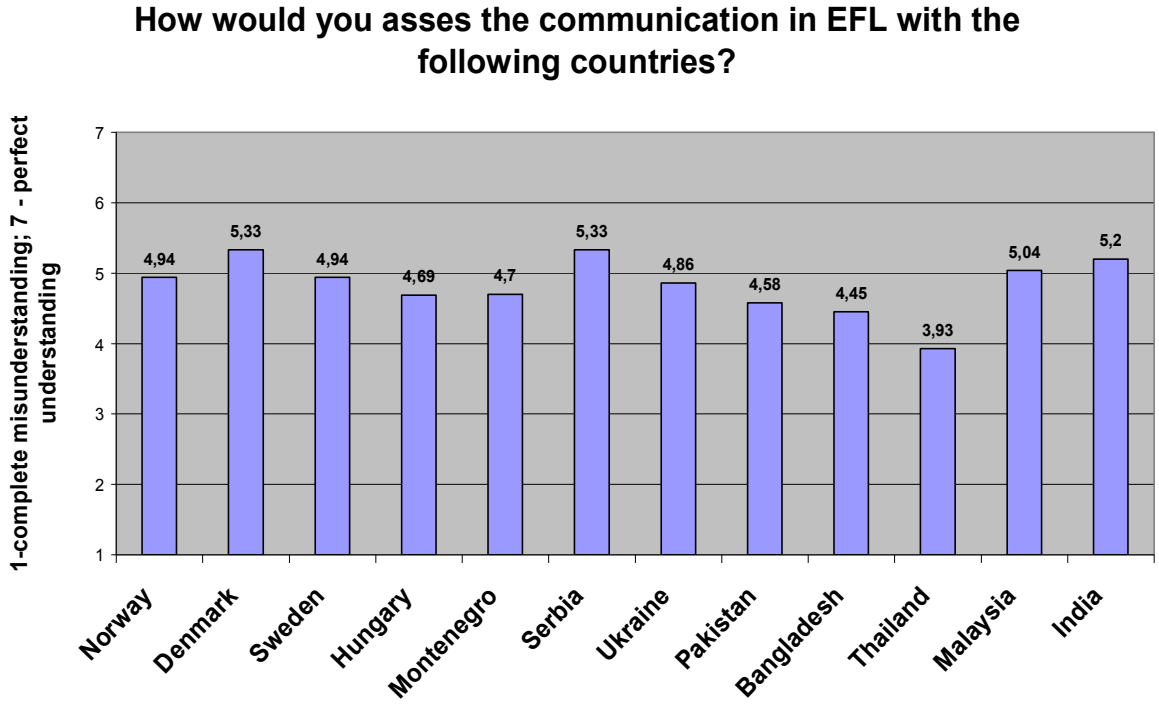


Figure 16: The communication in EFL with 12 countries; 1 is extremely low, 7 is extremely high

Two destinations which are assessed around the average level (4) for communication in EFL are Thailand and Bangladesh. With all the other countries the communication in EFL is assessed above the average level and hence could not be evaluated as problematic as a whole, but there are small exceptions, as it becomes clear from the interviews. Many of the respondents shared that they have troubles understanding accented speech. What they often do in order to overcome this is to ask for repetition with the excuse that the connection is not good:

"Blaming the technology is "easier". I have never addressed the problem to the accent, but to technology" shared a researcher form Telenor. The same informant adds that: *"The people has distinct accent and at the beginning it is difficult to understand what they are saying. But I am adapting to the accent or get used to it and it is definitely easier now."*

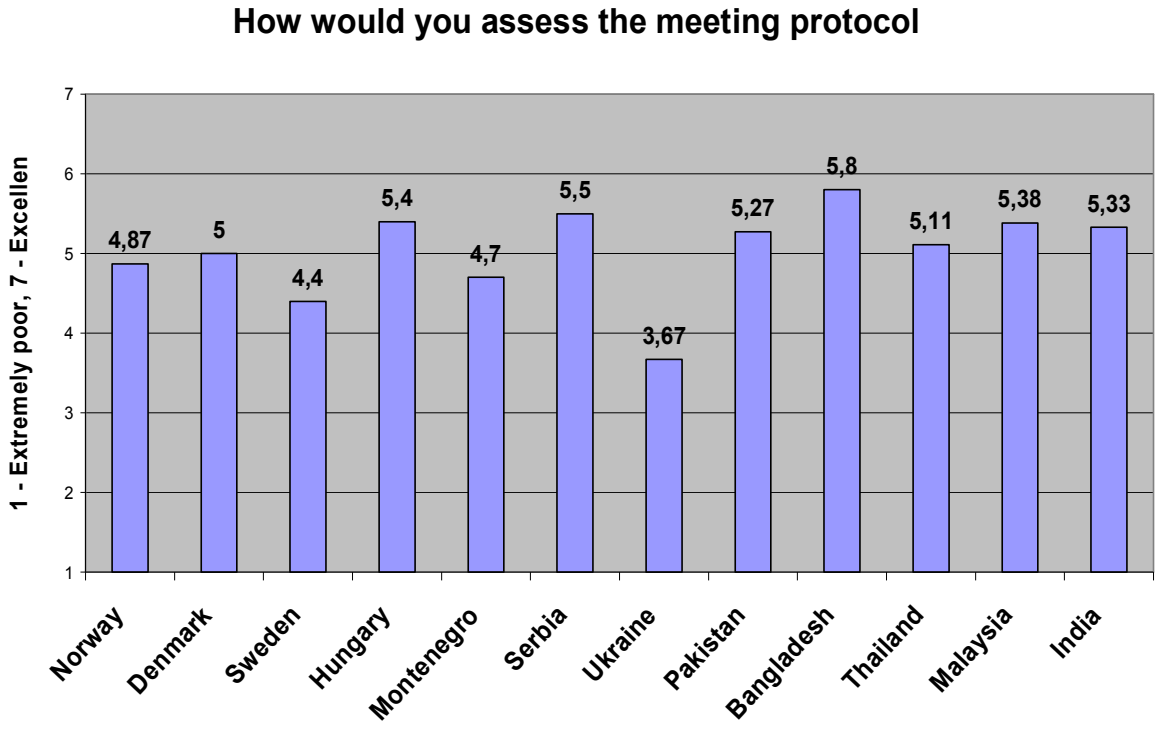


Figure 17: The meeting protocols for 12 countries

The meeting protocol is the rules that should be followed during the activity. Where these rules are not accepted, there are premises for tension between the communicating subgroups and hence lower outcome of the activity. The meeting protocol in the communication with Sweden and Ukraine is assessed closest to the average, but the rate is lower compared to the other destinations. These numbers confirm the data from the interviews that the most problematic destination concerning the meeting etiquette is Ukraine. But there is not much a user can do about someone's "not adequate" behaviour, other than telling them that the connection is bad: a common approach, identified from the interviews that the people from Telenor are using in order to inform their remote attendees for difficulties caused by them, due to accented pronunciation or behaviour.

8.4.4 Tool

Audio conferencing equipment

The tool itself could be seen as the core of the all technological problems. The table top conference phones are used by 75% of the respondents in a loudspeaker mode to participate in audio conference meetings. This system is used as an ordinary handset phone by 21% of the respondents. VoIP audio conferencing with PC and headset is used by 47% of the respondents. Through a mobile phone the audio conference meetings are attended by 47% of the respondents. From the literature reviewed it became clear that the mobile phones pick up static and introduce noise in the audio conferences. Thereby it is not recommended to use them for conference meetings.

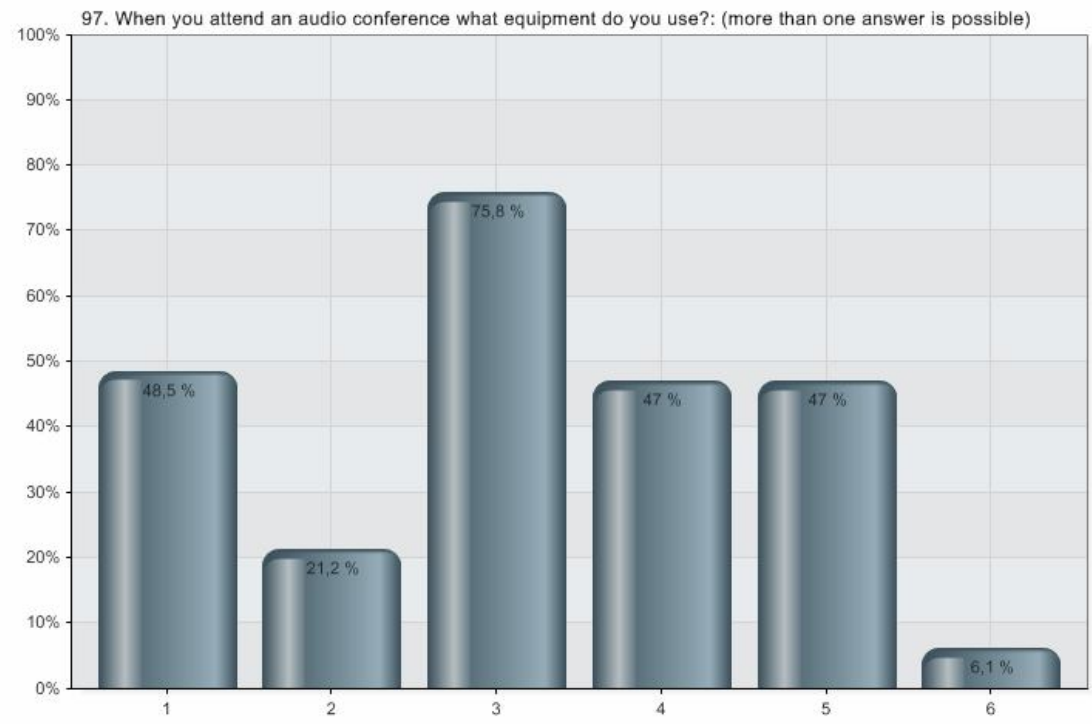


Figure 18: "What equipment do you use for AC? (Headset/ handset/ loudspeaker/ mobile phone/ PC/ other)"

Audio quality

92 % of the respondents have experienced problems with the audio quality. Almost 38 % of them have experienced it often and 61% occasionally. Noisy connection is

the problem most frequently experienced in AC. It is followed by echo, dropouts, speech clipping, delay, some parties could not be heard and voice distortion.

8.4.5 Relationship Subject – Tool

Using the systems

93.9% of the respondents argue that they know how to use audio conferencing. Most of them have learned to use it through the combination of the following three methods presented here by their frequency: their colleagues have showed them how to use it (50%), they intuitively have found out how it works (45.2%) and they have read the instructions (40.3%).

The rest 6.1% of the respondents have answered that they do not know how to use the system. The reason for it, according to them, is because half part of them are using it very rarely and the other half find it not user-friendly. The results from the personal interviews were approximately the same - only one of 12 respondents (which is 8,3 %) has answered that s/he could not use the audio conferencing systems in the meeting rooms, because they are not user friendly and s/he does not have access to instructions.

36% of the respondents know how to set up a multipoint conference call without reading the instructions. 47% need instructions in order to initiate a multipoint conference call, but since they are not available, as it became clear from the face-to-face interviews, it means probably that almost half part of the respondents could not manage to set up such a call. 16.7% could set up a multipoint conference call only through the use of bridge, where in practice all the calling parties are dialing the same bridge number in reserved time.

Selection of communication media combination

The results from the survey have shown that the audio conferencing is used by 95.5% of the respondents for project work/ meeting with colleagues . From the interviews it

became clear that at the beginning of the projects the researchers at Telenor prefer to meet face-to-face with their distributed colleagues in order to ensure shared understanding of the work topic(s) and *"global orientation which gives meaning to human processes"* (Heeren & Lewis, 1997, p.89). This level of work activity is described by Heeren & Lewis (1997) as *intentional level*. When the researchers have to take important decisions *"in order to achieve a final goal or intermediate goal"* (Heeren & Lewis, 1997, p.89) during the real work on the project then they are using both audio conferencing, video-conferencing or face-to-face meetings, but mostly audio conferencing, became clear from the interviews. This is the *functional level* which is oriented to planning and problem solving. The work on *operational level* consists of *"practical routines which are a prerequisite to the conscious, purposeful actions at the functional level"* (Heeren & Lewis, 1997, p.89). One such a routine in Telenor research department, as it became clear from the interviews, are the presentations in which some of the researchers have shared that they are participating almost every day and often the presentations last for about 1- 2 hours. During the presentations one or more participants are presenting the results of their work on operational level in order to inform the distributed participants in the project. A survey result shows that 49.3% of the respondents are using audio conferencing for presentations. These examples illustrate that the audio conferencing systems in Telenor seem to be used mainly on functional and operational level.

Intentional level

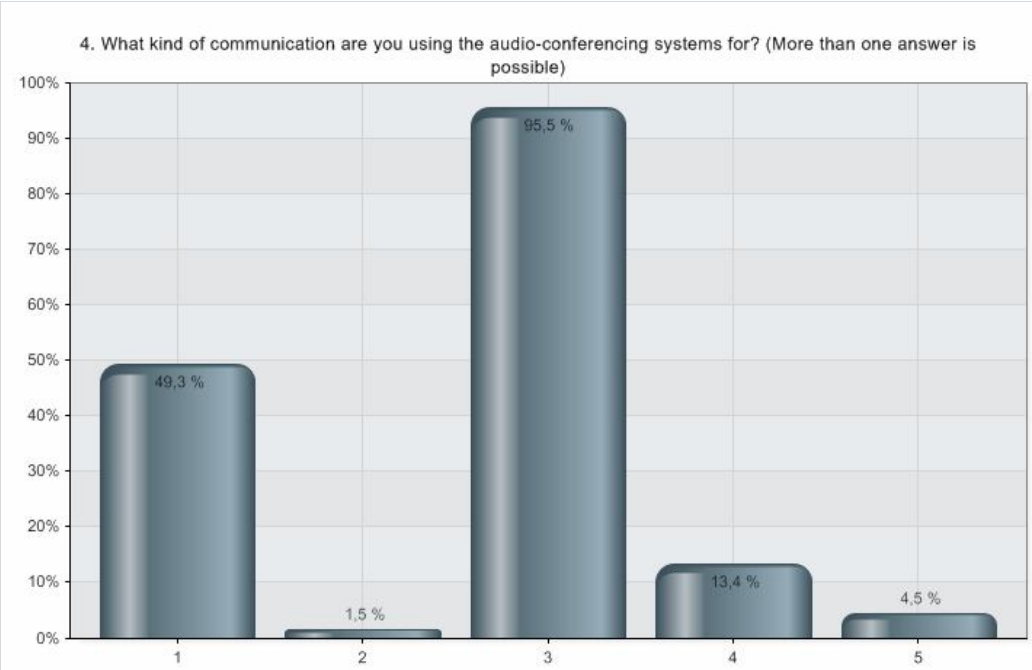
Functional level

Operational level

Figure 19: Hierarhical model of AT

Another result from is that 13.4% of the respondents are using audio conferencing for meetings with clients. In Telenor not all of the people are working directly with clients, most of them in the research department, for instance, are working mostly on projects with different colleagues' groups. In its essence the meetings with clients are

mainly focused on negotiations between the client and Telenor about the clients' needs and desires from one side and with the possibilities of the company to fulfill the client's needs and expectations. This type of work is on intentional level and requires a richer medium and minimum equivocality and is usually preferred to be mediated through face-to-face meetings.

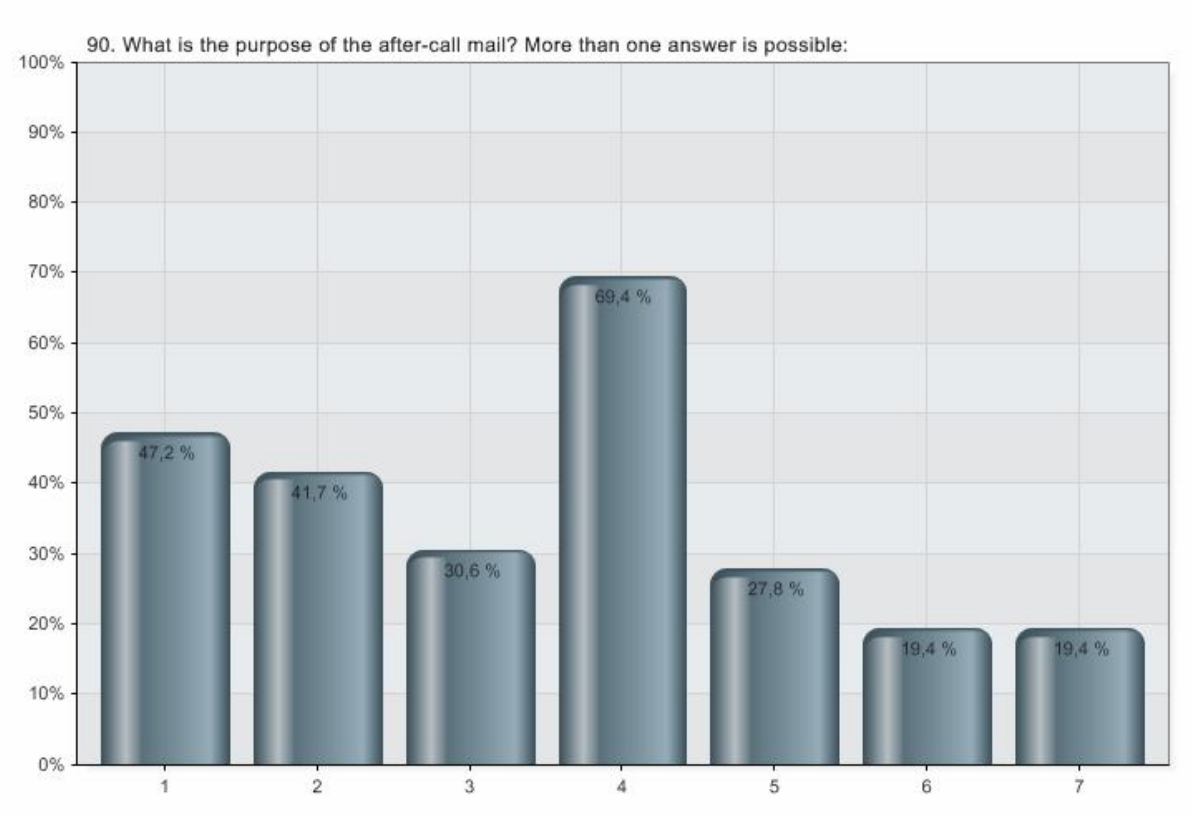


Alternatives		Percent	Value
1	Presentations	49,3 %	33
2	Social meetings	1,5 %	1
3	Project work/ meetings with colleagues	95,5 %	64
4	Meetings with clients	13,4 %	9
5	Other meetings, please specify:	4,5 %	3
Total			110

Figure 20: "What kind of communication are you using the audio conferencing for?"

An audio conference call is usually combined with an after-call email. 56% of the employees in R&I are using after-call mail in accordance with the specific of their work. When the respondents were asked "What is the purpose of the after-call mail"

() 67% of them has answered with “Common understanding”, 47% with “Clarity” and 41% with “Consensus”.



Clarity/ Consensus/ Conformation/ Common understanding/ Reinforcement/ “Legality” of agreement/ Other

Figure 21: “What is the purpose of the after-call mail?”

8.4.6 Relationship Subject – Rules

To the question “Are you aware of the existence of a specific set of conventions for audio conference meetings, so called meeting protocol?” 64.2% have answered with ‘No’ and 35.8% with ‘Yes’. The question includes an explanation about what is meant by ‘meeting protocol’: *Meeting protocol is a series of non-written rules and conventions which the participants of the meeting are following in order to have a better communication environment. Following a meeting protocol contributes for more effective and more efficient audio conference meetings. For example, people*

may agree to mute their microphone when not talking to improve the overall audio quality of the current speaker. (See Appendix D)

It seems that although a significantly big part of the responds do not know what the meeting protocol is, it appeared after the explanation, that they were applying it in practice, as it is evident from the following chart. To the question "Do you follow the meeting protocol?" 66.7% have answered positively and 33.3% negatively.

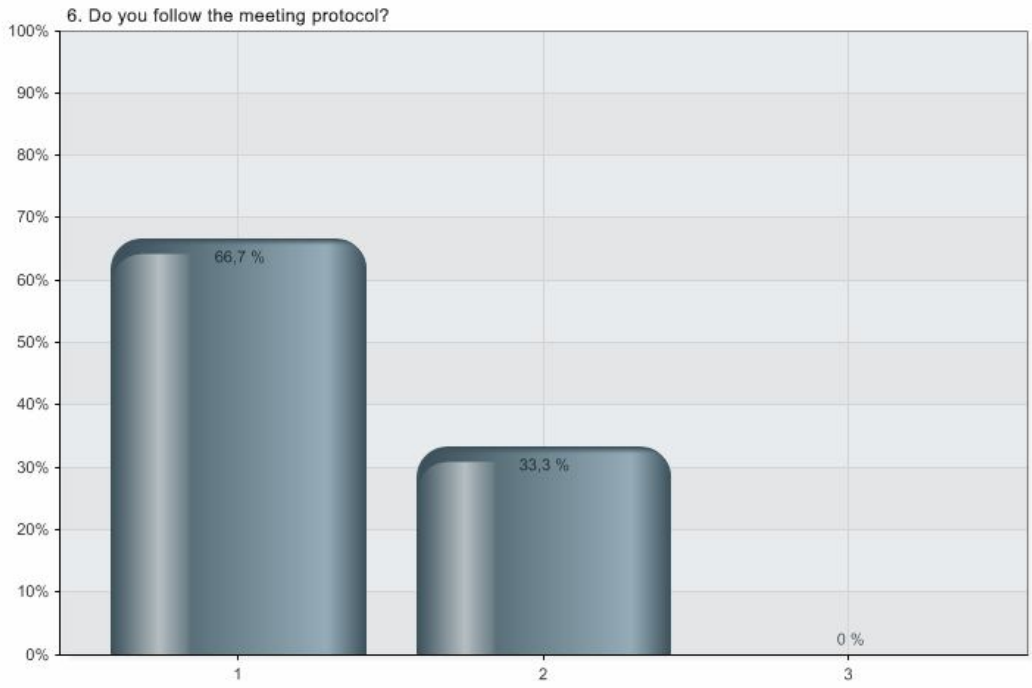


Figure 22:"Do you follow the meeting protocol?"

Some of the respondents were intrigued by the questions about the conference protocol / etiquette and have left such a comment:

"Tell us where we can find the "protocol" to use during audio conferences."

The users can use the list of guidelines for audio conferencing (see Appendix A) to improve the outcome of the activity and to enhance QoE. By sending these guidelines to their distributed colleagues or partners all the parties could agree to use this audio conferencing protocol / etiquette. This could solve the problem with those participants

in the audio conference meetings who involuntarily are creating communication difficulties, as for example background noise; in Telenor nobody has addressed this problem personally, because it could be accepted as rudeness.

Consciousness

Nardi (1996) explains that *"a basic tenet of activity theory is that a notion of consciousness is central to a depiction of activity. Vygotsky described consciousness as a phenomenon that unifies attention, intention, memory, reasoning, and speech..."* (Nardi, 1996). By applying the notion of consciousness to the audio conferencing as an activity, the subject's consciousness could be interpreted as adequate behavior during the activity. This following description is a try to quantify a qualitative notion, which supports the attempts for understanding the dynamics in the activity system: When the subject of the activity participates with their *"attention, intention, memory, reasoning, and speech"* (Nardi, 1996) then the outcome (or effectiveness - the overall performance) of the activity is more likely to be satisfactory, in comparison to when the level of the consciousness with which the subject participates in the activity is lower.

The subject's consciousness could be improved by support of the element Tool from the system. The series of works by Yankelovich (2004, 2005, 2006) illustrate the idea of engaging the tool (audio conferencing system called Meeting Central) in the role of providing means for consciousness (or *"attention, intention, memory, reasoning, and speech"* (Nardi, 1996)). *"The system provides visibility and awareness of who is in a meeting, who is speaking, who has joined or left, who is presenting, who is on mute, who is voting, and who is having audio problems"* (Yankelovich, 2004). In the next chapter of the current work it is presented an approach – Guidelines for effective audio conferencing/ Meeting protocol, which is engaging the Rules in the role of providing means for consciousness.

8.4.7 Community

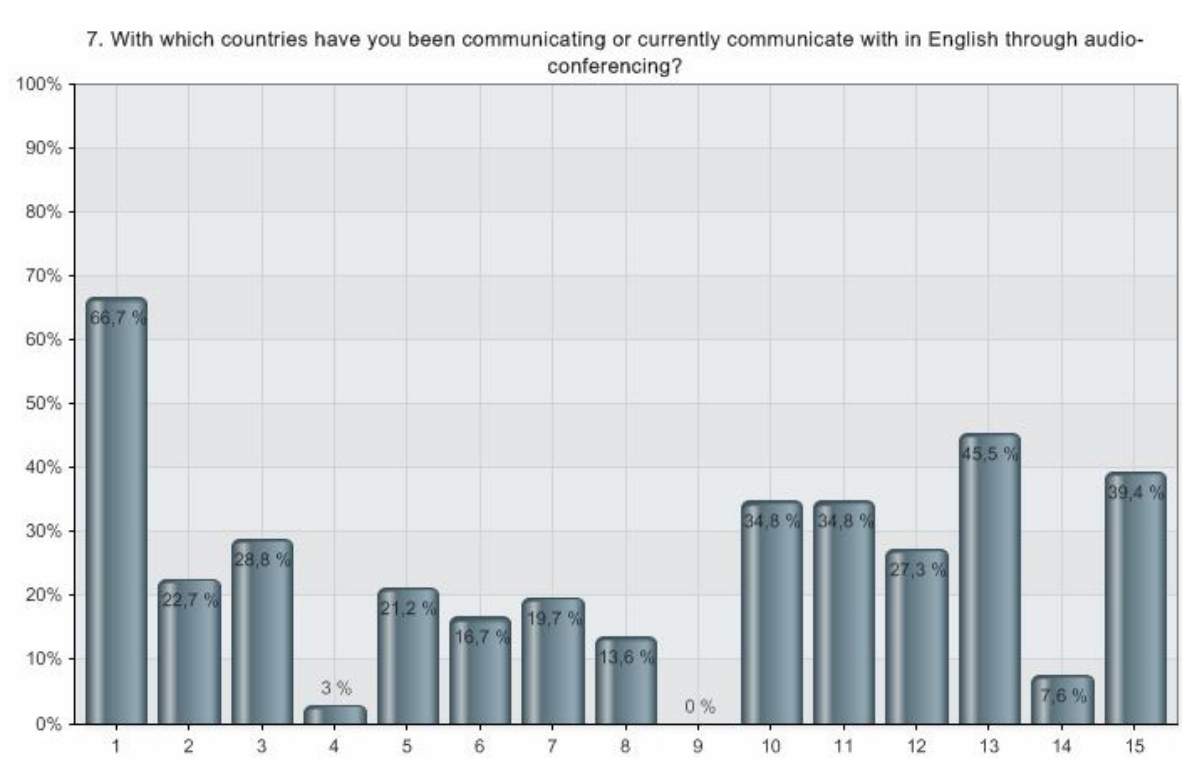


Figure 23:"With which countries have you been communicating or currently communicate with in English through audio conferencing?"

Norway - 66,7 %, Denmark - 22,7 %, Sweden - 28,8 %, Finland - 3,0 %, Hungary - 1,2 %, Montenegro - 16,7 %, Serbia - 19,7 %, Ukraine - 13,6 %, Russia - 0,0 %, Pakistan - 34,8 %, Bangladesh - 34,8 %, Thailand - 27,3 %, Malaysia - 45,5 %, India - 7,6 %, Other - 39,4 %

This chart represents the community of countries the respondents are communicating with. By using this distribution the management could direct its attention to introduce more effective and efficient means of communication with the most frequent destinations, like Norway, Malaysia, Bangladesh, Pakistan, Sweden, Thailand and Denmark. A similar decision concerning the community of the audio conferencing activity is proposed by a respondent to the survey, who has left his comment about the question 'How to improve the quality of audio conferencing":

"*WoW project need to finance proper connectivity between Telenor Locations, including local dial in numbers"*

The countries on the chart are all the locations where Telenor operates on their local market. The respondents working with those locations think that Telenor as a big, international, innovative telecommunication company has to have a proper connectivity between all the locations.

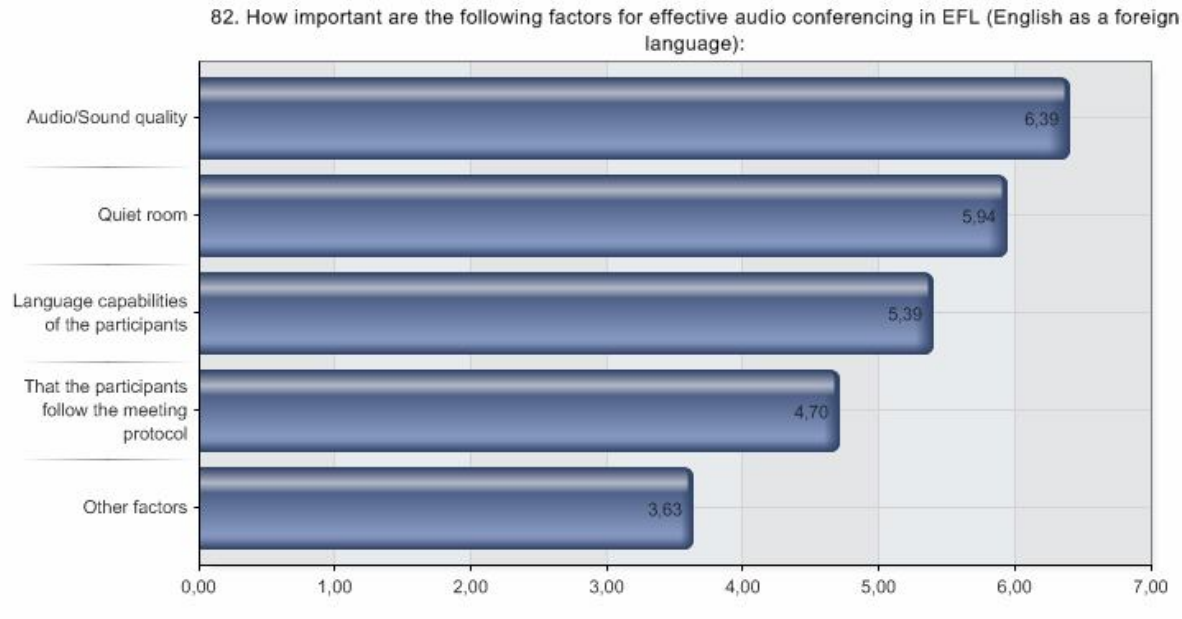


Figure 24: "How important are the following factors for effective audio conferencing in EFL?"

The respondents were asked to rate the importance of the factors audio/sound quality, quiet room, language capabilities of the participants, the participants following the meeting protocol for effective audio conferencing in EFL. They have answered that the audio quality is the most important issue (with a score 6.39) followed by "quiet room" (5.94), "language capabilities of the participants" (5.39) and "the participants are following the meeting protocol" (4.70). Although that the *"users are often unable to correctly identify what is responsible for the problem in a conferencing environment, due to the complexity of many interacting factors and media"* (Watson & Sasse, 2000) they are unanimous that the audio (speech) quality is the most important factor for effective audio conferencing. Most of the research is focused namely on this issue, for example (Yankelovich, 2006), (Rodman, 2006). These

results could be used as a starting point of the endeavors for improving the effectiveness of audio conferencing sessions in Telenor.

8.4.8 Outcome of activity

Efficiency, effectiveness and satisfaction

The outcome of the activity is measured subjectively by the participants. The three traditional usability metrics are efficiency, effectiveness and basic subjective satisfaction:

Efficiency is defined as *"resources expended in relation to the accuracy and completeness with which users achieve goals"* ISO 9241-11:1998(E). The statistical mean is 4.56. It is determined by the respondent's opinion using a scale from 1 to 7.

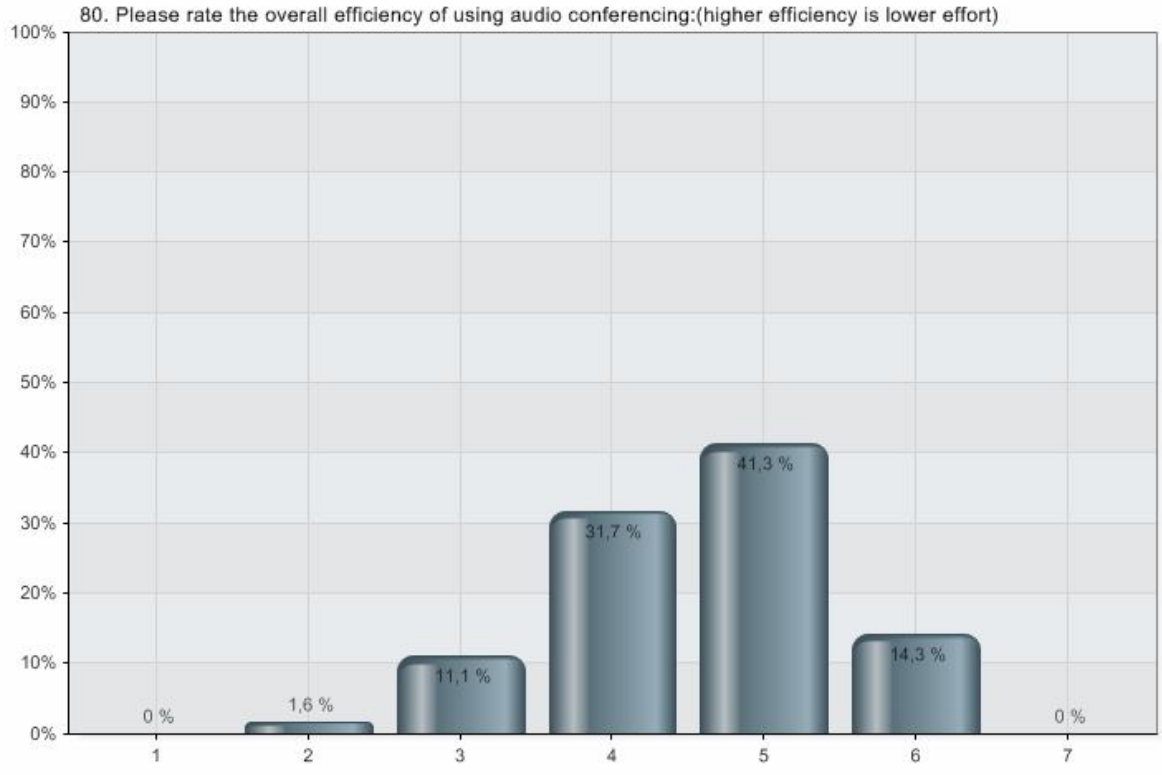


Figure 25: "Please rate the overall efficiency of using audio conferencing."

Effectiveness is *"accuracy and completeness with which users achieve specified goals"* ISO 9241-11:1998(E). The statistical mean is 4,69 using a scale from 1 to 7.

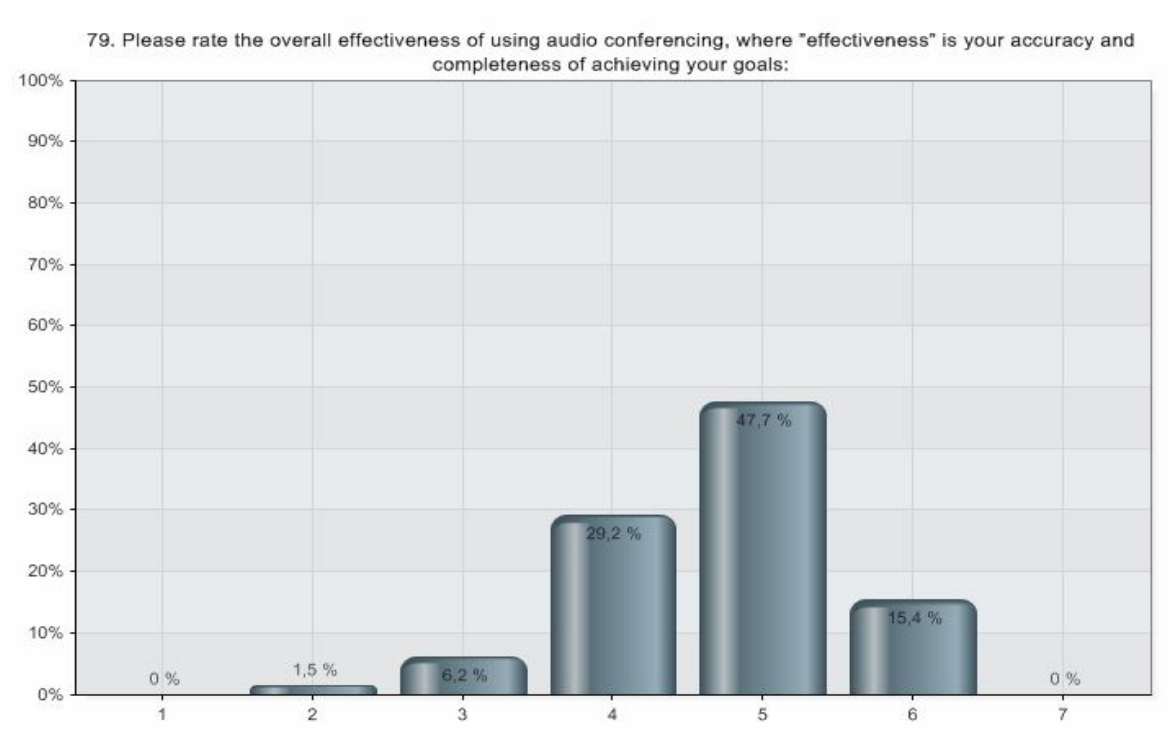


Figure 26: "Please rate the overall effectiveness of using audio conferencing."

Satisfaction is defined as *"freedom from discomfort, and positive attitudes towards the use of the product"* ISO 9241-11:1998(E). The statistical mean for satisfaction is 4,51 using a scale from 1 to 7.

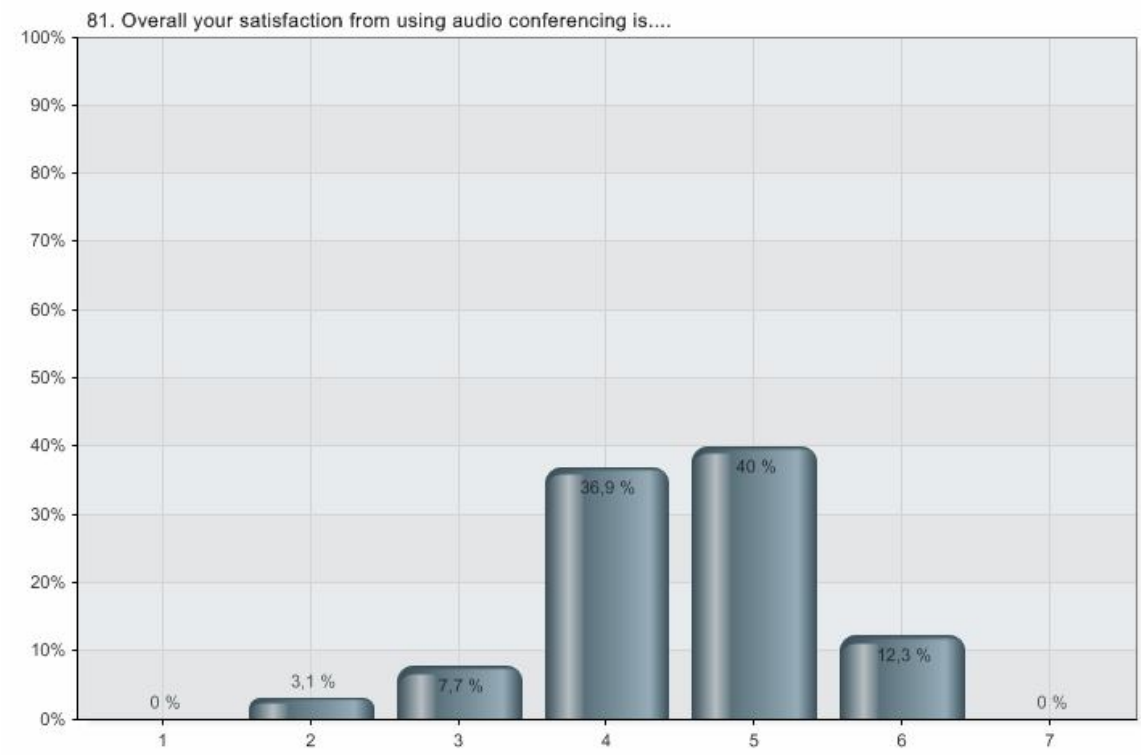


Figure 27: "Overall your satisfaction from using audio conferencing is:"

Improving the outcome of activity

Six respondents have shared their opinion about how to improve the outcomes of activity (audio conferencing). Although it became clear that the behavioural issues create the most serious difficulties in audio conferences, the users' rationality is searching for technical solutions of many of the problems:

"Encourage the use of proven commercially available services such as Skype rather than inferior home made solutions."

"Provide more (distributed) microphones in the audio conferencing rooms."

"Multiple microphones for long meeting tables, possibility for muting each mic separately."

"The equipment used should be consistent over meeting rooms"

"WoW project need to finance proper connectivity between Telenor Locations, including local dial in numbers"

"Improve the audio quality, Add possibilities for recording audio and video"

9. Conclusion

The current chapter synthesizes the findings from the interview, observation and questionnaire studies and builds a new understanding of the investigated phenomenon. All the results described in those studies provided thorough and detailed understanding of the relationships between the factors influencing the audio conferencing.

This work investigated what kind of limitations and restrictions are created by communicating in EFL when using the audio conferencing systems and services in Telenor. The three main domains that these limitations and restrictions belong to were identified as: – technical, socio-cultural and language based group of issues. These three groups supported the investigation part of revealing challenges in audio conferencing sessions and also the analytical part. For more detailed analysis the AT was used. AT supported the analysis in a way that the framework offered by the theory was used as a lens through which all the revealed issues from the investigation could be related to each other in order to be understood. In addition the hierarchical model of activity was used in order to understand how the audio conferencing supports the communication processes in the company – when, in which situations, and in what kind of communication the audio conferencing systems and services are preferred to be used.

As a first step, the research identified the weaknesses in the communication practices (when using audio conferencing systems in EFL) in Telenor.

The second step was to develop methods through which the gaps in the investigated spheres of communication could be filled. One such method is the guidelines for meeting etiquette. In addition the gap between subject and tool could be filled by additional training of the subjects in using the tools. It is not investigated and proven whether those methods will enhance the QoE.

These were the main research questions followed by the main findings from the research:

9.1 What are the technical difficulties and what kind of approaches could be offered to overcome them?

The technological perspective was examined from the level on which the user interacts with technology. It was revealed that the lack of knowledge about using the technique decreases drastically the quality of the call. Approximately half part of the users of these systems are using them as ordinary telephones, because they do not find it user friendly and easy-to-use and do not have instructions in order to learn how to use the different built-in features. About 7% of the employees could not either include more than one party in a conference call.

Another key factor for the success of audio conferencing is audio quality, because if the audio quality is poor then the primary objective of the call could be lost. Audio quality is in a strong relationship with the participants' behavior and could be improved or worsen by it. Thereby the developed guidelines are a solution for improving audio quality. In addition participants could control the audio quality by selecting the most proper equipment among various audio conferencing system and devices in the department. For example many of the new table top phones have integrated functions for noise and echo cancellations that could enhance QoE. Half part of the interviewees is using mobile phone for AC, which decreases the overall audio quality in the meeting. This issue could be overcome if the meeting participants are aware of the meeting guidelines.

A big part of the users suggested that the problems accompanying audio conferencing sessions could be solved by technology improvements. In order to fulfill the users' expectations the new and improved technology should not only have improved features but they have to be more user-friendly and more easy-to-learn.

9.2 What are the socio-cultural difficulties and what kind of approaches could be offered to overcome them?

The employees in Telenor R&I can choose from several different audio conferencing systems to use for conducting meeting sessions (see 4.2). The motives for those choices were discussed in the chapters "Selection of communication media" in 6.4.2 and 8.4.5. The quality of audio conferencing sessions can be influenced by factors with physical (network, terminal, meeting room characteristics) and nonphysical (behavior) characteristics. Despite the facts that the technicians can ensure a good system quality for a satisfactory session performance, the end users are those who by using the systems in a particular way are defining the end user experience. The end user behavior was found to be the factor that mostly decreases QoE. A major behavior problem is the late start of the meetings caused in some of the cases by the lack of knowledge of using the devices. Other important behavioral issues reducing QoE are talking over the conversation and attending the conference from public places.

Thereby it is argued that training users on technology and giving them guidelines on how to use technology and how to behave would improve the user experience in audio conferencing sessions. It was found that the need of following rules is greater when participants' number is bigger.

9.3 What are the linguistic difficulties and what kind of approaches could be offered to overcome them?

More than half part of the respondents has experienced problems in AC in EFL due to the language (page 91). The linguistic issue has a dual nature – technological and socio-cultural. The kinds of issues included in those two perspectives were analysed and described. From technological perspective the problems originate from the limited bandwidth. From a socio-cultural perspective it was revealed that the main linguistic obstacle is the accented speech. Despite this, the communication in EFL is

easier with people having similar first languages. Another important problem is the uncertainty to speak. This factor creates unbalanced communication network of subjects, where the most active participants are those who feel confident when using English. From a socio-cultural perspective the language issue could be solved by introducing and following the guidelines for effective meeting behavior.

9.4 Contributions

The current work developed three main valuable contributions: AS's matrix which is a theoretical based solution developed for analytical purposes; Quantitative study offers statistics that may be applied towards improving QoE, in all three domains: technological, socio-cultural and language based; The guidelines for improving user experience in audio conferencing session, which is the essential outcome of the work based on both empirical and theoretical findings could further help with number of identified issues.

9.4.1 AS's Matrix

A new analytical approach based on AT and using AS's elements has been developed. It was applied to the specific needs of the present investigation. The approach is simple and is represented by Table 1. It puts together the relationships between each two elements from AS with the aim to find the problematic relationships and further objective to explore how to bridge the gap between the two elements. This approach could be freely adopted by any investigation using the concepts of AT, because AT itself does not have a developed methodology or algorithm through which an investigation could be conducted. The approach offer simplicity for investigating and revealing gaps in any activity system.

9.4.2 Categorised statistical data for the current case

The communication difficulties with 12 different countries were evaluated through the results of the survey. The statistical data categorized by technical, socio-cultural and linguistic issues could be used by the managers for strategic planning and decision making and from the employees for orientation for future communication situations with a particular destination. It is assumed that the use of this information could enhance the quality of experience. A future investigation could be conducted in order to prove to what extend it holds true.

9.4.3 Guidelines for improving user experience: Meeting protocol / etiquette

The research extracted guidelines for effective audio conferencing behavior for the participants in the audio conferencing meetings. It was found that the need of introducing and using such rules is greater when there are more than 11 remote attendees in the meeting and when there are more than 3 or 4 locations. It is assumed that by offering this approach a number of technical, socio-cultural and linguistic difficulties in communication could be solved and the meeting effectiveness could be enhanced, but there have not been done any empirical experiments to prove this.

An implication for a future work is to investigate whether the users are willing to adopt such audio conferencing etiquette, how they do this and whether the outcomes of the adoption enhance the quality of user experience.

It is assumed that the developed guidelines for audio conferencing etiquette could be used by any user of audio conferencing in any setting.

10. Future work

Any statements in this paper are restricted only to the current domain of investigation. Whether the findings from this domain could be applied to a broader area of applications of audio conferencing in other domains would be the natural continuation of this work. There are assumptions that there are similarities with other activity systems where audio conferencing systems are used in other circumstances, but it should be investigated and proven to what extend it holds true. It could be done by applying the same research approach to another setting/s and compare the results.

The data collected in the study is only subjective – the user opinion. Objective data, such as network performance was not collected. In order to ensure better reliability of the subjective data it could be collected objective data for the following purposes:

- Parallel with user opinion data collection to be gathered data about the network performance. These two types of data could be matched in order to find the relationship between user behaviour and network performance. This approach is usually applied to works investigating the relationships between QoE and QoS, as in (Watson & Sasse, 1998).

- It could be conducted an empirical experiment in order to prove the hypothesis that the use of the introduced here meeting protocol will enhance the meeting effectiveness.

References

Aanestad, M. (2003). The Camera as an Actor: Design-in-Use of Telemedicine Infrastructure in Surgery, Computer Supported Cooperative Work: The Journal of Collaborative Computing, vol 12, pp 1-20.

Bardram, J. E. (1998). Designing for the Dynamics of Cooperative Work Activities, Poltrock & Grudin (red): Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW'98, ACM, pp. 89-98.

Barthelmess, P. and Anderson, K.M. (2002). A View of Software Development Environments Based on Activity Theory, *Computer Supperted Cooperative Work*, vol. 11, nos. 1–2, pp. 13–37.

Beaney, M. (2009). *Analysis,* http://plato.stanford.edu, (as 12 February 2010).

Berg, M. (1997). On Distribution, Drift and the Electronic Medical Record: Some Tools for a Sociology of the Formal,. Hughes m.fl. (red): Proceedings of the Fifth European Conference on Computer-Supported Cooperative Work, ECSCW'97, Kluwer, pp. 141-156 .

Blomberg, J., Giacomi, J. Mosher, A. and Swenton-Wall, P. (1993). Ethnographic Field Methods and Their Relation to Design, In Schuler, D. & A. Namioka (eds.) Participatory Design: Principles and Practices. Hillsdale, N.J.: Erlbaum. Pp. 123-156.

Bowers, J. (1994). The work to make a network work: studying CSCW in action, In Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work (Chapel Hill, North Carolina, United States, October 22 - 26, 1994). CSCW '94. ACM, New York, NY, 287-298.

Bryant, S., Forte A. and Bruckman A. (2005). Becoming Wikipedian: Transformation of participation in a collaborative online encyclopaedia, Proceedings of GROUP International Conference on Supporting Group Work, 2005. pp 1.-10

Button, G. (ed.) (1991). Ethnomethodology and the Human Sciences, Cambridge: Cambridge University Press.

Carstensen, P.H. and M. Nielsen (2000). Guiding the Thrust! Analytical Concepts in the Service of Coordination Support Systems, In R. Dieng, A. Giboin, L. Karsenty and G.D. Michelis (eds.), Designing Cooperative Systems: the Use of Theories and Models. Proceedings of the 5th International Conference on the Design of Cooperative Systems (COOP '2000). Amsterdam: IOS Press.

Chandler D. (1994). The Sapir-Whorf Hypothesis, UWA, http://www.aber.ac.uk, (as 22 June 2009).

Cogdill, S., Fanderclai, T. L., Kilborn, J., and Williams, M.G. (2001) Backchannel: Whispering in digital conversation, In Proc. of the 34th Hawaii International Conference on System Sciences (HICSS 2001); IEEE Press.

Conferencing, CommsBusiness, http://www.commsbusiness.co.uk, (as 22 April 2010).

Creative Research Systems, Survey Design, (2009), http://www.surveysystem.com, (as 28 Sep. 2009).

Daft, R. L., Lengel, R. H., and Trevino, L. K. (1987). Message equivocality, media selection and manager performance: implications for information systems. *MIS Q.* 11, 3 (Sep. 1987), 355-366.

Davis W. A A., Weinstein M. I. (2005). The Business Case for videoconferencing-Achieving a Competitive Edge, Wainhouse Research, March 2005, www.polycom.com, (as 08 February 2010).

Ember, C., Ember, M. (2006). Cultural Anthropology, Prentice Hall, Chapter One.

Engeström, Y. (1995). Innovative organizational learning in medical and legal settings, in Martin, L.M.W., Nelson, K., Tobach, E. (Eds),*Sociocultural Psychology: Theory and Practice of Doing and Knowing*, Cambridge University Press, Cambridge.

Engeström, Y. (1996). Developmental work research as educational research: looking ten years back and into the zone of proximal development, *Nordisk Pedagogik: Journal of Nordic Educational Research*, Vol. 16 pp.131-43.

Engeström, Y. (2001). Expansive Learning at Work: toward an activity theoretical reconceptualization, In Journal of Education and Work, Vol. 14, No. 1, 2001, pp. 133-156. Routledge.

Engeström, Y. and Nardi, B. (1998). A web on the wind: the structure of invisible work, Computer Supported Cooperative Work, volume 8, numbers 1-2 (special issue) at http://www.best.com/~/InvisibleW.html.

Erickson, T., Smith, D. N., Kellogg, W. A., Laff, M., Richards, J. T., and Bradner, E., (1999). Socially translucent systems: social proxies, persistent conversation, and the design of "babble". In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: the CHI Is the Limit* (Pittsburgh, Pennsylvania, United States, May 15 - 20, 1999). CHI '99. ACM, New York, NY, 72-79.

Erickson, T. and Kellogg, W. A. (2000). Social translucence: an approach to designing systems that support social processes, *ACM Trans. Comput.-Hum. Interact.* 7, 1 (Mar. 2000), 59-83.

ETSI Guide EG 202 670, STF354 Guidelines, http://portal.etsi.org, (as 17 Apr. 2010).

Fitzpatrick, G.; Tolone, W. J. & Kaplan, S. M. (1995). Work, Locales and Distributed Social Worlds, Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work, ECSCW'95, Kluwer, pp. 1-16 .

Forester J., Peters S., Hittleman M. (2005). Transcribing interviews, Cornell University, http://courses.cit.cornell.edu/practicestories/index.htm, (as 29 January 2010).

Forsythe, D. (1999). It's Just a Matter of Common Sense: Ethnography as Invisible Work, *CSCW* 8: 127-145.

de Freitas, M. R. and Byrne, E. (2006). Activity theory as an analytical tool: a case study of IS development for an anti-retroviral treatment clinic in South Africa. In *Proceedings of the 2006 Annual Research Conference of the South African institute of Computer Scientists and information Technologists on IT Research in Developing Countries* (Somerset West, South Africa, October 09 - 11, 2006). J. Bishop and D. Kourie, Eds. SAICSIT, vol. 204. South African Institute for Computer Scientists and Information Technologists.

Frohlich, D. and Luff P. (1989). Conversational Resources for Situated Action. *Proceedings of the SIGCHI Conference on Wings for the Mind*. Austin, TX: ACM, pp. 253–258.

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Engelwood Cliffs, NJ: Prentice-Hall.

Gasser, L. (1986). *The integration of computing and routine work*, ACM; ACM Transactions on Information Systems (TOIS), Volume 4 Issue 3.

Glaser, Barney G & Strauss, Anselm L. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research, Chicago, Aldine Publishing Company

Halverson, C. A. (2002). Activity Theory and Distributed Cognition: Or What Does CSCW Need to DO with Theories?. Comput. Supported Coop. Work 11, 1-2 (Apr. 2002), 243-267.

Hanseth, O. & Monteiro, E. (1997) Inscribing Behaviour in Information Infrastructure Standards,. Accounting, Management & Information Technology vol. 7 no. 4, pp. 183-211.

Hassan, M. and Alekseevich, D.F. (2006) Variable packet size of ip packets for voip transmission", In IASTED International Conference, conference on Internet and Multimedia Systems and Applications (IMSA), 136–141, ACTA Press.

Heeren, E. & Lewis, R. (1997) Selecting communication media for distributed communities, Journal of Computer Assisted Learning (1997) 13, 85-98.

Heim, J., Brooks, P., Følstad, A., Schleimann, T., Hestnes, B., Ulseth, T., Heiestad, S., Frowein, H., Devoldere, P., Aaby, C., O'Malley, C., Brundell, P., Londsdale, P., (2001) Final verification of real-time communication requirements, (Available from www.eye-to-eye.org).

Heim, J., Brooks, P., Følstad, A., Schliemann, T., Hestnes, B., Heiestad, S., et al. (2003) Fitness-for-Purpose Evaluation Methodology. (Available from www.eye-to-eye.org)

Herbsleb, J.D., Mockus, A., Finholt, T.A., & Grinter, R.E. (2000) Distance, Dependencies, and Delay in a Global Collaboration, In Proceedings, *ACM Conference on Computer-Supported Cooperative Work* (CSCW), pages 319-328, Philadelphia, PA, Dec. 2-7.

Hutchins, E.L. (1995b) Cognition in the Wild, Cambridge, MA: MIT Press.

Iraossi, G. (2006) The Power of Survey Design: A User's Guide for Managing Surveys, Interpreting Results, and Influencing Respondents", Washington, D.C.: The World Bank.

ISO 9241-11:1998(E), (available at http://www.iso.org).

ITU-T Recommendation E.800 (08/94), Terms and definitions related to quality of service and network performance including dependability.

ITU-T Recommendation G.1000 (11/2001), "Communications quality of service: A framework and definitions".

ITU-T, P.10/G.100, http://www.itu.int/md/, (as 09 September 2009)

Kaptelinin, V. & Nardi, B.A. (2006) Acting with technology: activity theory and interaction design, Cambridge, MA: MIT Press.

Katzenberg, B. and McDermott J. (1994) Meaning-making in the Creation of Useful SummaryReports, Proceedings of the Conference on Computer Supported Cooperative Work. Chapel Hill, NC: ACM, pp. 199–206.

Kim T., Chang A., Holland L., Pentland A. S. (2008) Meeting mediator: enhancing group collaboration using sociometric feedback, Proceedings of the ACM 2008 conference on Computer supported cooperative work, November 08-12, San Diego, CA, USA.

Koszalka, T. A., Wu, Chun-Ping (2004).A Cultural Historical Activity Theory [CHAT] Analysis of Technology Integration: Case Study of Two Teachers, Association for Educational Communications and Technology, 27th, Chicago, IL, October 19-23, 2004.

Kuutti, K. (1991) The Concept of Activity as a Basic Unit of Analysis for CSCW Research,. Bannon et al (red): Proceedings of the 2nd European Conference on Computer-Supported Cooperative Work, ECSCW'91, Kluwer, pp. 249-264.

Law, E., Roto, V., Hassenzahl, M., Vermeeren, A., Kort, J. (2009) Understanding, Scoping and Defining User Experience: A Survey Approach, In Proceedings of Human Factors in Computing Systems conference, CHI'09. April 4-9, 2009, Boston, MA, USA.

Modaff, J.V., & Modaff, D.P. (2000) Technical notes on audio recording, Research on Language and Social Interaction, 33(1), 101-118.

Muhammad N., Chiavelli, D., Soldani, D., and Li, M. (2006) Interoduction in Soldani, D., Li, M., Cuny, R., (eds.), QoS and QoE Management in UMTS Cellular Systems, John Wiley & Sons, 2006 .

Nardi, Bonnie A. (1996) Context and Consciousness: Activity Theory and Human-computer Interaction", pp 1-20. MIT Press, Cambridge, MA.

Nardi, B. (2000). An Ecological Approach to Design, http://www.nardi/best.com, (as 20 April 2010).

Nardi, B., (2005), http://darrouzet-nardi.net, (as 20 April 2010).

Nardi, B. A., & Redmiles, D. F. (2002). Activity Theory and the Practice of Design", Special Issue of Computer Supported Cooperative Work. *Computer-Supported Cooperative Work, 11*(1-2).

Nathanael, D., Marmaras, N., Papantoniou, B., Zarboutis, N. (2002) Socio-technical Systems Analysis: Which approach should be followed?, In Cognition, Culture & Design, S. Bagniara, S. Pozzi, A. Rizzo & P. Wright (Eds.), Sienna: Instututo di Sienze et Tecnologie dela Cognizione, pp. 137-142

O'Brien, M. (1993). Social research and sociology, in N. Gilbert (ed.), Researching social life, London: Sage, pp.1-17.

Orlikowski, W. (1992). Learning from Notes: Organizational Issues in Groupware Implementation, Turner and Kraut (eds.): Proceedings of the Conference on Computer-Supported Cooperative Work, CSCW'92, pp. 362-369.

Personopplysningsloven (Personal Data Act), Act of 14 April 2000 No. 31 relating to the processing of personal data**,** www.datatilsynet.no, (as 27 March 2010)

Ricardson, L. & St. Pierre, E., A. (2005) Writing: A Method of Inquiry. In Denzin, N. K. & Linkoln, Y. S. (eds.): *The Sage Handbook of Qualitative Research*. Sage Publications.

Rodman, J. (2006) The effect of bandwidth on speech intelligibility, Fellow/CTO, Polycom.

Sacks, H. (1992) *Lectures on Conversation*. Oxford: Blackwell.

Sacks, H., E. Schegloff and Jefferson G. (1978). A Simplest Systematics for the Organization of Turn-taking in Conversation. In J. Schenkein (ed.), *Studies in the Organization of Conversational Interaction*. New York, NY: Academic Press.

Sapir, E. (1929). 'The Status of Linguistics as a Science'. In E. Sapir (1958): Culture, Language and Personality (ed. D. G. Mandelbaum). Berkeley, CA: University of California Press

Schiff, N.D. and Purpura, K.P. (1997). The thalamic intralaminar nuclei: A role in visual awareness. The Neuroscientist 3, 8-15.

Schmidt, K. & L. Bannon: *Taking CSCW Seriously. Supporting Articulation Work*, (1992). Computer Supported Cooperative Work: The Journal of Collaborative Computing, vol. 1 no. 1, pp. 7-40.

Schmidt, K. and Simone C. (1996). Coordination Mechanisms: Towards a Conceptual Foundation of CSCW Systems Design. Computer Supported Cooperative Work, The Journal of Collaborative Computing, vol. 5, nos. 2–3, pp. 155–200.

Schofield J. (2008). *Can speech-recognition software transcribe interviews?*, Guardian, 21 February 2008, www.guardian.co.uk, (as 29 January 2010).

Shapiro, D. (1994). The Limits of Ethnography: Combining Social Sciences for CSCW. Computer Supported Cooperative Work, Chapel Hill, NC, ACM.

Shaw, M., (1976). Group dynamics: The psychology of small group behavior. McGraw-Hill.

Smart Survey, http://s3.amazonaws.com/SurveyMonkeyFiles/SmartSurvey.pdf, (as 28 Sep. 2009)

Strauss, A. (1978). A Social World Perspective. Pp. 119-28 in *Studies in Symbolic Interaction*, Volume 1, edited by Norman Denzin. Greenwich, CT: JAI Press.

Strauss, A. & Corbin, J., (1998). *Basics of qualitative research. Techniques and procedures for developing grounded theory*. Thousand Oaks, CA and London: Sage.

Suchman, L. A. (1987) *Plans and Situated Actions: the Problem of Human-Machine Communication*. Cambridge University Press.

Survey Design, http://www.surveysystem.com/sdesign.htm, (as 28 Sep. 2009)

The E-Model, E-Model tutorial, (2008), www.itu.int, (as 15 February 2010).

Watson, A. and Sasse, M. A. (1998) Measuring perceived quality of speech and video in multimedia conferencing applications, In *Proceedings of the Sixth ACM international Conference on Multimedia* (Bristol, United Kingdom, September 13 - 16, 1998). MULTIMEDIA '98. ACM, New York, NY, 55-60.

Watson, A. and Sasse, M. A. (2000) The good, the bad, and the muffled: the impact of different degradations on Internet speech. In *Proceedings of the Eighth ACM international Conference on Multimedia* (Marina del Rey, California, United States). MULTIMEDIA '00. ACM, New York, NY, 269-276.

Wikipedia, *Ethnography*, http://en.wikipedia.org/wiki/Ethnography, (as of Apr. 19, 2010, 16:42 GMT).

Wikipedia, *Public switched telephone network*, http://en.wikipedia.org/wiki/Public_switched_telephone_network, (as of Apr. 19, 2010, 16:24 GMT).

Wikipedia, *Telenor*, http://en.wikipedia.org/wiki/Telenor, (as of Apr. 19, 2010, 16:36 GMT).

Wikipedia, *User experience*, http://en.wikipedia.org/wiki/User_experience, (as of Apr. 19, 2010, 16:32 GMT).

Yankelovich, N., Kaplan, J., Provino, J., Wessler, M., and DiMicco, J. M., (2006) Improving audio conferencing: are two ears better than one?, In *Proceedings of the*

2006 20th Anniversary Conference on Computer Supported Cooperative Work (Banff, Alberta, Canada, November 04 - 08, 2006). CSCW '06. ACM, New York, NY, 333-342.

Appendix A: Guidelines for improving user experience in audio conferencing sessions: meeting protocol / etiquette

Socio-cultural issues

Be on time

You have to be ready for the meeting 5-10 min before the scheduled time. If you go 2 minutes before the start of the meeting there is a high probability that you will be late for the actual start of the meeting. You have to provide time for testing the equipment and preparing the room for the meeting (chairs, doors, windows, etc.).

Ensure that you are situated in a quiet environment

Close the doors, windows, switch off the mobile phone. The quiet environment contributes to improving the overall effectiveness of the audio conference meeting. The background noise could reduce the audio quality and could introduce disturbances that reflect to the participants' concentration and the meeting schedule.

Introduce yourself in the begingin of the meeting

By doing this your colleagues can link your voice with your name and the impression that you create in the beginning. They could more easily recognise your voice when you speak during the meeting.

Introduce yourself before speaking

By doing this your colleagues will not be in doubt who is speaking and will not lose time for asking additional question for clarifying who has said what.

Avoid typing on keyboard

Avoid typing on keyboard when you are close to the microphone. Could you imagine that your colleagues hear the produced sound several times louder than the sound you hear, because of the microphone!

Avoid whispering

Avoid whispering to your colleagues during an audio conference meeting. You could think that nobody hears you, but actually all the participants in the meeting are able to hear the sound, more loudly than you expect, because of the microphone. It disturbs the participants' concentration and reduces the meeting effectiveness.

Keep to the schedule

Remain consistent with the schedule in order to respect the time of the other participants.

Avoid multi-tasking

Avoid multi-tasking (talking on another phone, working on computer, typing sms) during a conference meeting.

Language issues

Speak slowly and attentively in the begining

In the beginning of the meeting you could afford to speak slowly in order to help your colleagues to get used to your English pronunciation – accent or dialect. After that you could switch to the normal tempo if you are sure that the distributed participants in the meeting do not have problems understanding your way of speaking English.

Ensure that your partners understand you

Use backchannel expressions to 'show' that you understand or do not understand. Linguistic backchannel utterances such as "uh-huh" or "yes, quite" indicate that you are listening and agree with what is being said.

After-call email

An approach for improving the effectiveness of audio conferencing sessions is the use of the combination of audio conference call and after-call email. The after-call email ensures common understanding, clarity and consensus.

Technical issues (on users level)

Technical issues in terms of how the user can control the technique.

Test the equipment

Test the technical condition of the equipment before important meetings.

Avoid the use of mobile phones

Avoid attending the conference by using mobile phones because they are likely to pick up static (noise, distortion of the broadcast signal caused by the electromagnetic radiation in the atmosphere).

Use the mute button

When you and your colleagues in the room are not engaged actively in a discussion it is better that you mute your phone / device. It can vastly contribute to improving the overall quality of the audio conference meeting - first by preventing the background noise to reduce the audio quality and second by preventing from disruptions that could minimise the concentration of the participants and to disorder the meeting schedule.

Appendix B: General Concepts and Terminology

Source: ITU-T Recommendation E800

quality of service (QoS)

The collective effect of service performance which determine the degree of satisfaction of a *user* of the *service*.

NOTES

1 The *quality of service* is characterized by the combined aspects of *service support performance*, *service operability performance*, *serveability performance*, *service security performance* and other factors specific to each service.

2 The term “quality of service” is not used to express a degree of excellence in a comparative sense nor is it used in a quantitative sense for technical evaluations. In these cases a qualifying adjective (modifier) should be used.

network performance

The ability of a network or network portion to provide the functions related to *communications* between *users*.

NOTES

1 Network performance applies to the Network Provider’s planning, development, operations and maintenance and is the detailed technical part of QOS, excluding service support performance and human factors.

2 Network performance is the main influence on serveability performance.

3 Network performance measures are meaningful to network providers and are quantifiable at the part of the network to which they apply. Quality of service measures are only quantifiable at a service access point.

4 It is up to the Network Provider to combine the Network Performance parameters in such a way that the economic requirements of the Network Provider, as well as the satisfaction of the User, are both fulfilled.

Service

A set of functions offered to a user by an organization.

User

Any entity external to the network which utilizes connections through the network for communication.

Connection

An association of resources providing means for communication between two or more devices in, or attached to, a telecommunication network.

Appendix C: Notes taken from the second observation

These are the notes taken during the second meeting:

15:42 – there was a disturbing background noise, like someone was opening a door;

15:47 – for approximately 10-20 sec there was a distortion of the voice of one of the participants who was talking at that moment;

15:52 – the voice fade away for 5 to 10 sec;

15:53 – distortion of the voice, metallic sound;

15:55 - distortion of the voice, metallic sound;

16:12 – there is an electrical background noise, maybe from a device nearby to some of the microphones; it is often called that the devices (or a microphone) pick ups static electricity or just static.

16:16 – the same electrical background noise / static;

16:19 – the same very loud electrical background noise;

Appendix D: Survey Questionnaire

Audio conferencing: quality of experience questionnaire

The purpose of this survey is to explore the quality of experience for the users of the audio conferencing systems at Telenor. Conferencing system implies more that 2 calling parties. The survey aims at finding any problems related to audio conferencing, with a special focus on carrying out conferencing session in English as a second language. Many factors can contribute to reduced understanding during the conversation which is carried out in a foreign language, from technical to non-technical, such as the lack of conversational conventions.

The questionnaire is completely anonymous. The results are to be used purely for research purposes. This questionnaire will take 30 – 40 minutes.

Language

1. What is your first language?

   ☐English

   ☐Norwegian

   ☐Other. Please specify!

2. Which other languages do you speak?

   ☐English

   ☐Norwegian

   ☐Other. Please specify!

3. Do you use English at work?

   3.1. No

   3.2. Yes

      3.2.1. How often?

      ☐ Daily/ every day

      ☐ A few times per week

      ☐ Occasionally (a few times per month)

      ☐ Rarely

Communication

Purpose and language

4. For what kind of communication are you using the audio-conferencing system?

   More than one answer is possible.

   4.1. Presentations

   4.2. Social meetings

   4.3. Discussions

   4.4. Project work/ meetings with colleagues

   4.5. Meetings with clients

   4.6. Other. Please specify!

Meeting Protocol

Meeting protocol is a series of non-written rules and conventions which the participants of the meeting are following in order to have a better communication environment. Following a meeting protocol contributes for more effective and more efficient audio conference meetings.

5. Are you aware of the existence of a specific set of conventions for audio conference meetings, so called meeting protocol?

   5.1. No

   5.2. Yes

      5.2.1. Do you follow the meeting protocol?

      ☐ Yes

      ☐ Sometimes

      ☐ No

Destination

6. With which countries have you been communicating or currently communicate with in English through audio-conferencing?

   6.1. Norway

      6.1.1. How would you assess the audio quality:

      ☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

      Extremely low Extremely high It varies

6.1.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.1.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely poor Excellent It varies

6.1.4. Specify which aspects of the communication are not satisfying?

6.2. Denmark

6.2.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.2.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.2.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.2.4. Specify which aspects of the communication are not satisfying?

6.3. Sweden

6.3.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.3.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.3.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.3.4. Specify which aspects of the communication are not satisfying?

6.4. Finland

6.4.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.4.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.4.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.4.4. Specify which aspects of the communication are not satisfying?

6.5. Hungary

6.5.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.5.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.5.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.5.4. Specify which aspects of the communication are not satisfying?

6.6. Montenegro

6.6.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.6.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.6.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.6.4. Specify which aspects of the communication are not satisfying?

6.7. Serbia

6.7.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.7.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.7.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.7.4. Specify which aspects of the communication are not satisfying?

6.8. Ukraine

6.8.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.8.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.8.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.8.4. Specify which aspects of the communication are not satisfying?

6.9. Russia

6.9.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.9.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.9.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.9.4. Specify which aspects of the communication are not satisfying?

6.10. Pakistan

6.10.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.10.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.10.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.10.4. Specify which aspects of the communication are not satisfying?

6.11. Bangladesh

6.11.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.11.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.11.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.11.4. Specify which aspects of the communication are not satisfying?

6.12. Thailand

6.12.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.12.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.12.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.12.4. Specify which aspects of the communication are not satisfying?

6.13. Malaysia

6.13.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.13.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.13.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.13.4. Specify which aspects of the communication are not satisfying?

6.14. India

6.14.1. How would you assess the audio quality:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.14.2. How would you assess the communication in EFL (English as a foreign language):

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Complete misunderstanding Perfect understanding It varies

6.14.3. How would you assess the meeting protocol:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8

Extremely low Extremely high It varies

6.14.4. Specify which aspects of the communication are not satisfying?

Language

6.15. Have you communicated with anyone whose English you can not understand?

6.15.1. No

6.15.2. Yes

6.15.2.1. How often have you experienced this?

☐ Often

☐ Occasionally. Please specify!

☐ Rarely

☐ Other. Please specify!

6.15.2.2. From which country or countries were they?

Audio conferencing system

7. Do you know how to use the audio conferencing systems?

7.1. Yes

7.1.1. How did you learn to use the system? More than one answer is possible.

☐ My colleagues showed me how to use it

☐ I read the instructions

☐ Intuitively I tried to find out how it works

☐ I learned through observing colleagues

☐ Other. Please specify!

7.2. No

7.2.1. Why?

☐ It is not user friendly

☐ Other. Please specify!

8. Do you know how to set up a multipoint conference call?

☐ No, I can just call the bridge number

☐ Yes, if I read the instructions

☐ Yes, without reading the instructions

8.1.1. Please describe the sequence of actions with a few keywords!

9. How often are you using an audio-conferencing system at Telenor?

☐ Daily

☐ Weekly

☐ Monthly

☐ More rarely

☐ Other. Please specify!

Challenges in Audio conferencing

10. Could you describe the nature of the problems that you have most often experienced during audio-conferencing? More than one answer is possible:

☐ Mainly technical problems. Please explain what these are…

☐ Mainly language problems. Please explain what these are…

☐ Mainly socio-cultural communication problems. Please explain what these are…

☐ Other. Please specify!

11. Please rate the overall **effectiveness** of using audio conferencing, where ”effectiveness” is your accuracy and completeness of achieving your goals:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Extremely low **Extremely high**

12. Please rate the overall **efficiency** of using audio conferencing, where ”efficiency” is the effort required for you to achieve your goals:

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Extremely low **Extremely high**

13. Overall your **satisfaction** from using audio conferencing is....

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Extremely low **Extremely high**

14. How important are the following factors for effective audio conferencing in EFL (English as a foreign language):

14.1. Audio/Sound quality

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Extremely unimportant **Extremely important**

14.2. Quiet room

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Extremely unimportant **Extremely important**

14.3. Language capabilities of the participants

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Extremely not important **Extremely important**

14.4. That the participants follow the meeting protocol

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Extremely not important **Extremely important**

14.5. If other, please specify!

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Extremely not important **Extremely important**

Audio quality issues

15. How would you assess the **audio quality** of the audio conferencing system in Telenor?

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Extremely low **Extremely high**

16. Have you ever experienced problems with the audio quality?

16.1. No

16.2. Yes

16.2.1. How often have you experienced problems with the audio quality?

16.2.1.1. Always

16.2.1.2. Often. Specify how often!

16.2.1.3. Occasionally. Specify how often!

16.2.1.4. Never

16.2.2. How often have you experienced the following audio quality problems?

16.2.2.1. Echo

☐ Always

☐ Occasionally. Specify!

☐ Rarely

☐ Never

16.2.2.2. Delay

☐ Always

☐ Occasionally. Specify!

☐ Rarely

☐ Never

16.2.2.3. Dropouts

☐ Always

☐ Occasionally. Specify!

☐ Rarely

☐ Never

16.2.2.4. Speech clipping/ Words cut off

☐ Always

☐ Occasionally. Specify!

☐ Rarely

☐ Never

16.2.2.5. Noisy connection

☐ Always

☐ Occasionally. Specify!

☐ Rarely

☐ Never

16.2.2.6. Voice distortion/ Robotic sound

☐ Always

☐ Occasionally. Specify!

☐ Rarely

☐ Never

16.2.2.7. Some of the parties can not be heard

☐ Always

☐ Occasionally. Specify!

☐ Rarely

☐ Never

16.2.2.8. Other. Please specify!

☐ Always

☐ Occasionally. Specify!

☐ Rarely

☐ Never

16.2.3. Order the problems that you have experienced by their frequency! Start with the most often:

☐Echo

☐Delay

☐Dropouts

☐Speech clipping

☐Noisy connection

☐Others. Please specify!

16.2.4. When audio quality is poor do you do something to help continue the conversation?

Socio-cultural issues

17. Have you ever experienced connections with background noise?

    17.1. No

    17.2. Yes

        17.2.1. Please, describe the nature of the noise

18. Do you use after-call email?

    18.1. No

    18.2. Yes

        18.2.1. What is the purpose of the after-call mail? More than one answer is possible:

        ☐ Clarity

        ☐ Consensus

        ☐ Conformation

        ☐ Common understanding

        ☐ Reinforcement

        ☐ "Legality" of agreement

        ☐ Other. Specify!

19. Have you experienced that audio conferences start later than planned?

    19.1. No

    19.2. Yes

19.2.1. Define in minutes what for you is a “late start” for audio conferencing:

☐ Between 1 and 5 minutes after the appointed time

☐ Between 5 and 10 minutes after the appointed time

☐ Between 10 and 15 minutes after the appointed time

☐ More than 15 minutes after the appointed time

19.2.2. Describe in what context does the late start of audio conferencing happen:

19.2.2.1. Presentations

☐ Often

☐ Occasionally

☐ Rarely

☐ Never

☐ I do not know

19.2.2.2. Social meetings

☐ Often

☐ Occasionally

☐ Rarely

☐Never

☐I do not know

19.2.2.3. Discussions

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not know

19.2.2.4. Monitoring and managing work processes

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not know

19.2.2.5. Project work/ meetings with colleagues

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not know

19.2.2.6. Meetings with clients

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not know

19.2.2.7. Other. Please specify!

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not know

19.2.3. Describe what are the most often reasons for late start of audio conferencing:

☐Technical difficulties in setting up the conference

☐People coming late for the meeting

☐Misunderstanding

☐Other. Please describe!

20. Have you ever felt removed or isolated in an audio-conference meeting?

20.1. No

20.2. Yes

20.2.1. In which situations have you felt removed or isolated in an audio-conference meeting?

21. Do you think you have ever been not understood during an audio conference?

21.1. No

21.2. Yes

21.2.1. Do you think you know why?

21.2.1.1. No

21.2.1.2. Yes

☐You spoke fast

☐You described a complex (or technical) subject

☐Other. Please specify!

22. Do you do anything in order to make your voice recognizable when you speak during the conference?

22.1. No

22.2. Yes

22.2.1. Please specify!

23. Please rate the following list of behaviors, the ones that you feel distract or reduce the quality of your audio conferencing meetings. Fill / write down other examples of such behavior, if any:

23.1. Moving doors, windows, chairs

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.1.1. How often have you experienced the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.2. Hearing the electrical sound from the near-by devices

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.2.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.3. Attending the conference from (or walking in) public places

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.3.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.4. Typing on the computer

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.4.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.5. Speaking on another (mobile) phone during a meeting

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.5.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.6. Whispering with colleagues

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.6.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.7. Coming late to the meeting

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.7.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.8. Overlapping /"talking over" the conversation

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.8.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.9. Breathing in the microphone

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.9.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.10. Interrupting

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.10.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.11. You do not understand her/his accented/dialect way of speaking English

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.11.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.12. She/ He can not understand your English

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.12.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.13. Drinking or eating

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.13.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.14. Speaking too fast

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.14.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.15. Speaking too slowly

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.15.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.16. Speaking too loudly

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.16.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.17. Speaking too quietly/ softly

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.17.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

23.18. Have you been a witness of any other patterns of destructive behavior that you want to add?

23.18.1. No

23.18.2. Yes

23.18.2.1. Please specify

☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7

Not destructive **Extremely destructive**

23.18.2.1.1. How often have you been witness to the aforementioned behavior:

☐Often

☐Occasionally

☐Rarely

☐Never

☐I do not remember

24. When you attend an audio conference what equipment do you use?: (more than one answer is possible)

24.1. Mobile phone

24.1.1. Why would you choose the mobile phone?

☐Mobility

☐Better quality

☐Other. Please specify!

24.1.2. In which situations do you use mobile phone for attending audio conferencing?

24.2. Headset

24.2.1. Why?

24.3. Handset

24.3.1. Why?

24.4. Loudspeaker

24.4.1. Why?

24.5. Other. Please specify!

25. Do you use the “Mute” button during audio conference meetings?

25.1. Yes, always

25.1.1. Why?

25.2. Yes, sometimes

25.2.1. Why?

25.3. No

25.3.1. Why?

26. Do you usually introduce yourself when speaking for the first time?

26.1. Yes

26.1.1. Why?

26.2. No

26.2.1. Why?

Final questions

27. Is there anything that you do not know, but you would like to know as a user of the audio conferencing system?

27.1. No

27.2. Yes

27.2.1. Please specify!

28. Do you have any other recommendations for improving the quality of audio conferencing?

28.1. No

28.2. Yes

28.2.1. Please describe!

29. Are you willing to leave contact information for possible future interviews?

29.1. No

29.2. Yes

29.2.1. Fill your email address, your name, your phone number

Thank you for your help!